



Hiller, T. C. (2017). Peer Effects in Endogenous Networks. *Games and Economic Behavior*, 105, 349-367.
<https://doi.org/10.1016/j.geb.2017.07.010>

Peer reviewed version

License (if available):
CC BY-NC-ND

Link to published version (if available):
[10.1016/j.geb.2017.07.010](https://doi.org/10.1016/j.geb.2017.07.010)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Elsevier at <http://www.sciencedirect.com/science/article/pii/S0899825617301367>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Peer Effects in Endogenous Networks

Timo Hiller*

August 24, 2017

Abstract

This paper presents a model of strategic network formation with local complementarities in effort levels and positive local externalities. Results are obtained for a general class of payoff functions, which subsumes the linear-quadratic specification frequently used in theoretical and applied work. We assume homogeneous agents and characterize equilibria for two-sided and one-sided link formation. (Pairwise) Nash equilibrium networks are nested split graphs, which are a strict subset of core-periphery networks. We highlight the relevance of the convexity of the value function for obtaining these structures. More central agents are shown to exert more effort and obtain higher gross payoffs in equilibrium. However, net of linking cost, central agents may obtain strictly lower net payoffs. The curvature of the value function is also important for efficiency considerations. These findings are relevant for many social and economic phenomena, such as educational attainment, criminal activity, labor market participation, and R&D expenditures of firms.

Key Words: Strategic network formation, peer effects, strategic complements, positive externalities. **JEL Codes:** D62, D85.

*This paper is based on a chapter of my Ph.D. dissertation. I am grateful to my supervisors Fernando Vega-Redondo and Massimo Morelli for their support and guidance. I thank the editor, David C. Parkes, an anonymous associate editor and an anonymous referee for helpful comments, which have significantly improved the paper. I also thank Oliver Bätz, Sebastian Bervoets, Francis Bloch, Yann Bramoullé, Sergio Currarini, Antonio Cabrales, Frédéric Deroian, Jon Eguia, Andrea Galeotti, Sanjeev Goyal, Seungwon (Eugene) Jeong, Burkhard Krenn, In-Uck Park, Paolo Pin, Gonçalo Pina, Brian Rogers, Tim Schmidt-Eisenlohr, Siyang Xiong, Yanos Zylberberg and seminar participants of the Networks Working Group at the European University Institute, the UECE Lisbon Meetings 2010, the CTN 2013, SAET 2015, the University of Cambridge and Aix-Marseille University for valuable comments. All remaining errors are mine. Contact: timo.hiller@bristol.ac.uk, Address: Department of Economics, University of Bristol, 8 Woodland Road, BS8 1TN, UK.

1 Introduction

Peer effects and social structure play an important role in determining individual behavior and aggregate outcomes in many social and economic environments. This has been documented by a large body of empirical work, which finds that peer effects and network position are crucial for educational attainment, criminal activity, labor market participation and R&D expenditures of firms. In these settings an agent's optimal action and payoff is thought to depend directly on the action or payoff of others (peer effects), while the relevant reference group is determined by a network of bilateral relationships (social structure). Given the wide set of social and economic situations where peer effects and network structure are important for determining outcomes, we aim to understand why certain structures commonly arise and how these affect payoffs and incentives to exert effort. We therefore endogenize the network in the presence of peer effects, where, in accordance with empirical work, peer effects are assumed to induce positive local externalities and strategic complementarities.¹

The interpretation of a connection and an agent's effort level will depend on the particular application one has in mind. When considering bilateral R&D partnerships among firms, for example, then a R&D partnership can be thought of as a link in a network, while a firm's production level may be thought of as effort. Assuming that a firm's marginal production cost is decreasing in the production level of its R&D partners (e.g. due to spillovers from learning-by-doing), then production/effort levels of directly connected firms can be shown to display strategic complements and to induce positive externalities. We formally derive the commonly used linear-quadratic payoff specification for such an application in the next section. In contrast, in the case of crime, social interaction is considered to be an important channel facilitating individual criminal behavior. In this context a link is taken to be the social tie between two individuals, while an agent's criminal activity is interpreted as effort.²

Our simple model captures frequently observed patterns regarding effort levels, performance and network structure. Due to the reinforcing interplay between linking decisions and incentives to exert effort, ex-ante identical agents may obtain different ex-post outcomes. More central agents display higher effort levels and higher gross payoffs. This is much in line with the empirical literature.³ Perhaps surprisingly, however, more central agents may obtain strictly lower net payoffs (net of linking cost) in equilibrium. This is informative for empirical work, which generally disregards linking cost. Equilibrium networks are shown to be nested split graphs, which are a special case of core-periphery networks. Core-periphery structures are often observed in empirical studies and empirical support for nested split graphs has also been established in environments where peer effects and social structure are important.^{4,5} Note that, while nested split graphs are well known in the mathematical graph theory literature, these networks have only been identified recently in economics.⁶

¹See Hoxby (2000), Sacerdote (2001), Glaeser et al. (1996) and Case and Katz (1991) for educational attainment, Ludwig et al. (2001), Bayer et al. (2009) and Damm and Dustmann (2014) for crime and juvenile delinquency, Topa (2001) and Conley and Topa (2001) for labor market participation and Cohen and Levinthal (1989, 1990) and Levin and Reiss (1988) for R&D expenditures.

²For additional motivation, applications and derivations of the linear-quadratic payoff specification see König et al. (2014).

³See Baldwin et al. (1997), Lin (2001) Powell et al. (1996) and Ahuja (2000).

⁴See Adamic and Adam (2003), Canter (2004), Powell et al. (1996) and Baker et al. (2004).

⁵See Åkerman and Larsson (2014), Uzzi (1996) and Soramaki et al. (2007). König et al. (2014) provide a discussion of nested split graphs with further references, empirical evidence and an interpretation of the theoretical model in terms of networks of banks and trade networks.

⁶Goyal and Joshi (2003) is a very early paper that features nested split graphs (the authors call them

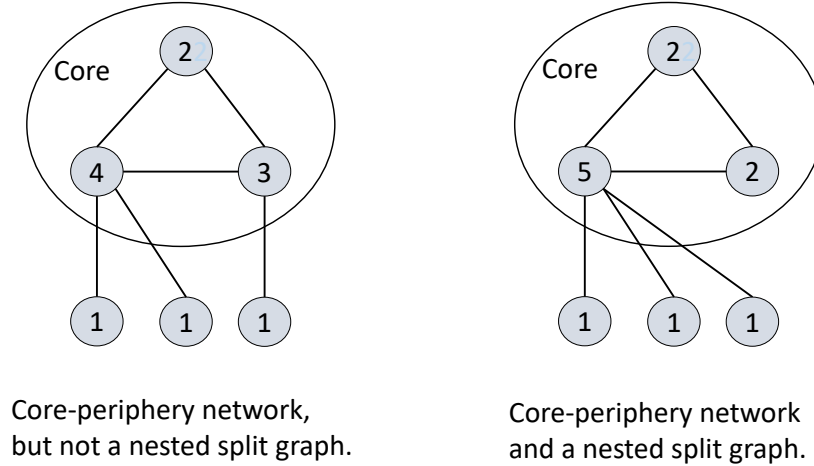


Figure 1: Core-periphery networks and nested split graphs

We illustrate the difference between core-periphery and nested split graphs in Figure 1 above, while formal definitions are provided in the Analysis section. In a core-periphery network the set of agents can be partitioned into two sets, called the core and the periphery, such that all agents in the core are connected to each other, while no pair of agents in the periphery is connected (Bramoullé, 2007). In a nested split graph the neighborhood of every agent is contained in the neighborhood of agents with a higher number of links.

The setup of our model is simple. Agents simultaneously choose a non-negative, continuous effort level and create links at a cost. The significance of a link is that direct neighbors in the network “access” or benefit from each other’s effort levels (positive local externalities) and an agent’s incentive to exert effort is increasing in the sum of direct neighbors’ effort levels (strategic complementarities). We assume payoff functions such that best response functions are either linear or, different from much of the prevailing literature, concave. Furthermore, the corresponding value function is assumed to be convex. That is, when best responding, own payoffs are convex in the sum of effort levels of direct neighbors. The economic interpretation of a convex value function is simple in our context. An agent’s incentives to create links to other agents, thereby benefiting from their exerted effort, are increasing in the effort level of agents to whom a connection is already established. We will show that this is what drives most of our results and, more specifically, provides a rationalization for the frequently observed core-periphery networks and nested split graphs. Note that a convex value function is consistent with payoff functions that are concave in the sum of neighbors’ effort levels (i.e. diminishing marginal returns) and may be interpreted as a strong form of strategic complementarity. Note further that in all the applications considered above, link formation is typically thought to be two-sided. This motivates the use of pairwise Nash equilibrium as our notion of strategic stability, since it reflects the bilateral nature of creating a link (and paying its cost).⁷ To the best of my knowledge our paper is the first to solve a two-sided network formation model in this context.

The paper’s contribution and its relation to the literature on networks is described toward the end of the paper and we only briefly summarize some of the remaining results here. Regarding the multiplicity of equilibria, we provide simple conditions such that a network,

interlinked stars).

⁷Pairwise Nash equilibrium was first discussed in Jackson and Wolinsky (1996). For applications see, for example, Goyal and Joshi (2003) and Belleflamme and Bloch (2004).

that may be thought of as intermediate relative to two pairwise Nash equilibrium networks, is also a pairwise Nash equilibrium network. Furthermore, we show that our results are robust to the introduction of small heterogeneity in payoff functions and linking cost. We also analyze the problem of a planner, who chooses agents' effort levels and may construct a network at a convex cost per link. Under a simple additional assumption on payoffs, which allows for the linear-quadratic payoff specification, all efficient networks are again shown to be a nested split graphs. Furthermore, we solve for the case of one-sided network formation and show that all Nash equilibrium networks are again nested split graphs.

The rest of the paper is organized as follows. Section 2 describes the model. In Section 3 we first present a result on the existence and uniqueness of a Nash equilibrium on a fixed network and then solve the model for two-sided network formation. Moreover, we study robustness and present our efficiency result. Section 4 describes the contribution of the paper and its relation to the networks literature. Section 5 provides a brief discussion and Section 5 concludes. In Appendix A conditions on payoffs are presented such that the best response function and value function fit our setup. The proofs for the two-sided model specification are relegated to Appendix B. The online appendix presents the case of one-sided network formation. All numerical calculations were executed in Mathematica and are available upon request.

2 Model Description

Let $N = \{1, 2, \dots, n\}$ be the set of players with $n \geq 3$. Each agent i chooses an effort level $x_i \in X$ and announces a set of agents to whom the agent wishes to be linked to, which is represented by a row vector $\mathbf{g}_i = (g_{i,1}, \dots, g_{i,i-1}, g_{i,i+1}, \dots, g_{i,n-1})$, with $g_{i,j} \in \{0, 1\}$ for each $j \in N \setminus \{i\}$. Assume $X = [0, +\infty)$ and $\mathbf{g}_i \in G_i = \{0, 1\}^{n-1}$. The set of agent i 's strategies is denoted by $S_i = X \times G_i$ and the set of strategies of all players by $S = S_1 \times S_2 \times \dots \times S_n$. A strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g}) \in S$ then specifies the individual effort level for each player, $\mathbf{x} = (x_1, x_2, \dots, x_n)$, and the set of intended links, $\mathbf{g} = (\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n)$. A link between i and j , denoted by $\bar{g}_{i,j} = 1$, is created if and only if *both* agents i and j intend to create a link. That is, $\bar{g}_{i,j} = 1$ if and only if $g_{i,j} = g_{j,i} = 1$ (and $\bar{g}_{i,j} = 0$ otherwise) and therefore $\bar{g}_{i,j} = \bar{g}_{j,i}$. We define the undirected graph $\bar{\mathbf{g}}$ as $\bar{\mathbf{g}} = \{\{i, j\} \in N : \bar{g}_{i,j} = 1\}$. That is, $\bar{\mathbf{g}}$ is a collection of links, which are listed as subsets of N of size 2. We write $\bar{\mathbf{g}} \subset \bar{\mathbf{g}}'$ to indicate that $\{\{i, j\} \in N : \{i, j\} \in \bar{\mathbf{g}}\} \subset \{\{i, j\} \in N : \{i, j\} \in \bar{\mathbf{g}}'\}$. The presence of a link $\bar{g}_{i,j} = 1$ allows players to directly benefit from the effort level exerted by the respective other agent involved in the link. Denote the set of i 's neighbors in $\bar{\mathbf{g}}$ with $N_i(\bar{\mathbf{g}}) = \{j \in N : \bar{g}_{i,j} = 1\}$ and the corresponding cardinality with $\eta_i(\bar{\mathbf{g}}) = |N_i(\bar{\mathbf{g}})|$.⁸ The aggregate effort level of agent i 's neighbors in $\bar{\mathbf{g}}$, i.e. the effort level accessed, is written as $y_i(\bar{\mathbf{g}}) = \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$. We sometimes write y_i for $y_i(\bar{\mathbf{g}})$ when it is clear from the context. Given a network $\bar{\mathbf{g}}$, $\bar{\mathbf{g}} + \bar{g}_{i,j}$ and $\bar{\mathbf{g}} - \bar{g}_{i,j}$ have the following interpretation. When $\bar{g}_{i,j} = 0$ in $\bar{\mathbf{g}}$, then $\bar{\mathbf{g}} + \bar{g}_{i,j}$ adds the link $\bar{g}_{i,j} = 1$, while if $\bar{g}_{i,j} = 1$ in $\bar{\mathbf{g}}$, then $\bar{\mathbf{g}} + \bar{g}_{i,j} = \bar{\mathbf{g}}$. Similarly, if $\bar{g}_{i,j} = 1$ in $\bar{\mathbf{g}}$, then $\bar{\mathbf{g}} - \bar{g}_{i,j}$ deletes the link $\bar{g}_{i,j}$, while if $\bar{g}_{i,j} = 0$ in $\bar{\mathbf{g}}$, then $\bar{\mathbf{g}} - \bar{g}_{i,j} = \bar{\mathbf{g}}$. The network is called empty and denoted by $\bar{\mathbf{g}}^e$ if $\bar{g}_{i,j} = 0 \forall i, j \in N$ and complete and denoted by $\bar{\mathbf{g}}^c$ if $\bar{g}_{i,j} = 1 \forall i, j \in N$.

⁸Note that we assume that agents are not linked to themselves and agents are not included in their own neighborhood.

Payoffs of player i under strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ are given by

$$\Pi_i(\mathbf{s}) = \pi_i(\mathbf{x}, \bar{\mathbf{g}}) - \eta_i(\bar{\mathbf{g}})\kappa,$$

where κ denotes linking cost with $\kappa > 0$. Gross payoffs, $\pi_i(\mathbf{x}, \bar{\mathbf{g}})$, are assumed to be a function of own effort, x_i , and the sum of effort levels of direct neighbors, $y_i(\bar{\mathbf{g}}) = \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$. For ease of notation we sometimes write $\pi_i(x_i, y_i)$ and drop the subscripts when they are clear from the context. We assume strict positive local externalities and strict strategic complementarities, so that $\partial\pi(x, y)/\partial y > 0$ and $\partial^2\pi(x, y)/\partial x\partial y > 0 \forall x, y$. Assumption 1 and Assumption 2 define conditions on the best response function and the value function directly. For corresponding properties on the payoff function see Appendix A.

Assumption 1 (Best response function). The unique best response of player i to the vector of effort levels \mathbf{x}_{-i} in network $\bar{\mathbf{g}}$ is given by

$$\bar{x}_i(\mathbf{x}_{-i}, \bar{\mathbf{g}}) = \bar{x}(\sum_{j \in N_i(\bar{\mathbf{g}})} x_j) \forall i \in N,$$

where $\bar{x}(0) > 0$, $0 \leq \lim_{y \rightarrow \infty} \bar{x}'(y) < \frac{1}{n-1}$ and either $\bar{x}''(y) < 0 \forall y$, or $\bar{x}''(y) = 0 \forall y$.

Assumption 2 (Value function). The maximized gross payoff under activity \mathbf{x}_{-i} in network $\bar{\mathbf{g}}$ is given by

$$\pi_i(\bar{x}_i, \mathbf{x}_{-i}, \bar{\mathbf{g}}) = v(\sum_{j \in N_i(\bar{\mathbf{g}})} x_j) \forall i \in N,$$

where $v(0) \geq 0$, $v'(y) > 0$ and $v''(y) \geq 0 \forall y$.⁹

Next, we briefly discuss the assumption of the convexity of the value function and relate it to properties of the payoff function. Taking the second derivative of $v(y)$ and applying the envelope theorem we can write

$$v''(y) = \left(\frac{\partial^2 \pi(x, y)}{\partial x \partial y} \Big|_{x=\bar{x}(y)} \times \frac{d\bar{x}(y)}{dy} \right) + \frac{\partial^2 \pi(x, y)}{\partial^2 y} \Big|_{x=\bar{x}(y)}.$$

Due to strict strategic complementarities, the first term in the above expression is positive and a sufficient condition for the value function to be convex is therefore that the second term is weakly greater than zero. Note that a convex value function is consistent with diminishing marginal returns in the sum of neighbors' effort levels. That is, the value function may be convex even if the second term is strictly smaller than zero. Below we present a payoff function with this property. The expression for the convexity of the value function allows us to interpret convex value functions in the presence of decreasing marginal returns, which may be thought of as a strong form of complementarity. The value function is convex if the nature of strategic interaction is strong and the following condition holds

$$\frac{\partial^2 \pi(x, y)}{\partial x \partial y} \Big|_{x=\bar{x}(y)} \times \frac{d\bar{x}(y)}{dy} \geq \left| \frac{\partial^2 \pi(x, y)}{\partial^2 y} \Big|_{x=\bar{x}(y)} \right| \forall x, y.¹⁰$$

⁹Note that $v'(y) > 0$ follows directly from positive externalities.

¹⁰Note that with strategic substitutes $\partial^2 \pi(x, y)/\partial x \partial y < 0$ and $d\bar{x}(y)/dy < 0$, so that in this case $\partial^2 \pi(x, y)/\partial x \partial y \Big|_{x=\bar{x}(y)} \times d\bar{x}(y)/dy$ is also positive. Therefore, in the case of strategic substitutes and decreasing marginal returns, a convex value function may be interpreted as a strong form of strategic substitutes.

Next we provide two examples of gross payoff functions that satisfy our assumptions. First we derive the frequently studied linear-quadratic payoff function in the context of a very simple model of R&D.¹¹ Assume there are n firms, which operate in n independent markets (i.e. each firm is a monopolist). Assume further that each firm i faces inverse demand function $p_i = \alpha - x_i$, where x_i is firm i 's production level. Firm i 's marginal cost function is given by $c_i(\bar{\mathbf{g}}) = \gamma - \lambda \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$. The interpretation of a link $\bar{g}_{i,j} = 1$ is a R&D agreement between firms i and j , which decreases marginal production cost (due to, for example, learning-by-doing effects). More specifically, assume firm i 's marginal production cost decreases linearly in the sum of production levels of its neighbors in $\bar{\mathbf{g}}$. Assume further that $\alpha > \gamma$ and that $\gamma > 0$ is sufficiently large, so that $c_i(\bar{\mathbf{g}}) > 0$. Gross payoffs are then given by $\pi_i(\mathbf{x}, \bar{\mathbf{g}}) = p_i x_i - c_i(\bar{\mathbf{g}}) x_i$, which can be rewritten as

$$\pi_i(\mathbf{x}, \bar{\mathbf{g}}) = (\alpha - \gamma) x_i - x_i^2 + \lambda x_i \sum_{j \in N_i(\bar{\mathbf{g}})} x_j.$$

Note that, after a normalization, this yields the classic linear-quadratic payoff function, which has been widely used to study peer effects (see, for example, Calvó-Armengol et al., 2005 and 2009 and König et al., 2014) and is presented in Example 1 below.

Example 1 Linear-quadratic payoffs:

$$\pi_i(\mathbf{x}, \bar{\mathbf{g}}) = x_i - \frac{\beta}{2} x_i^2 + \lambda x_i \sum_{j \in N_i(\bar{\mathbf{g}})} x_j,$$

where $\beta > 0$ and $\lambda > 0$. The corresponding best response and value function are given by:

$$\bar{x}_i = \frac{1 + \lambda \sum_{j \in N_i(\bar{\mathbf{g}})} x_j}{\beta} \text{ and } v_i = \frac{(1 + \lambda \sum_{j \in N_i(\bar{\mathbf{g}})} x_j)^2}{2\beta}.$$

Example 2 illustrates that our assumptions allow for payoff functions with diminishing marginal returns in the sum of neighbors' effort levels.¹² That is, diminishing marginal returns are compatible with a (strictly) convex value function.

Example 2 Payoffs with diminishing marginal returns and a (strictly) convex value function:

$$\pi_i(\mathbf{x}, \bar{\mathbf{g}}) = (1 + \sum_{j \in N_i(\bar{\mathbf{g}})} x_j)^q x_i - \frac{c}{2} x_i^2,$$

where $q \in (\frac{1}{2}, 1)$ and $c > 0$. The corresponding best response and value function are given by:

$$\bar{x}_i = \frac{1}{c} (1 + \sum_{j \in N_i(\bar{\mathbf{g}})} x_j)^q \text{ and } v_i = \frac{1}{2c} (1 + \sum_{j \in N_i(\bar{\mathbf{g}})} x_j)^{2q}.$$

Next we define *pairwise Nash equilibrium (PNE)* in the presence of simultaneous moves and effort choice. We assume that when agents i and j deviate to create a link, then deviation effort levels are mutual best responses (while the remaining agent's effort levels are assumed to remain unchanged). The corresponding deviation effort levels are denoted by $x'_i = \bar{x}(y_i(\bar{\mathbf{g}}) + x'_j)$.

¹¹For formal derivations of the linear-quadratic specification in the context of interbank lending, crime and trade, see König et al. (2014).

¹²Note that for our assumptions to hold we neither need to assume the separability of the utility function, nor a quadratic cost function.

A strategy profile $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$ is a pairwise Nash equilibrium *iff*

- for any $i \in N$ and every $\mathbf{s}_i \in S_i$, $\Pi_i(\mathbf{s}^*) \geq \Pi_i(\mathbf{s}_i, \mathbf{s}_{-i}^*)$;
- for all $\bar{g}_{i,j} = 0$, if $\Pi_i(x'_i, x'_j, \mathbf{x}_{-i,-j}^*, \bar{\mathbf{g}}^* + \bar{g}_{i,j}) > \Pi_i(\mathbf{s}^*)$,
then $\Pi_j(x'_i, x'_j, \mathbf{x}_{-i,-j}^*, \bar{\mathbf{g}}^* + \bar{g}_{i,j}) < \Pi_j(\mathbf{s}^*)$.

A pairwise Nash equilibrium is both a Nash equilibrium and pairwise stable and therefore refines Nash equilibrium. Pairwise Nash equilibrium allows for deviations where a pair of agents creates a link (and we assume that deviating agents best respond to each other's effort level) Furthermore, pairwise Nash equilibrium allows for deviations in which an agent deletes any subset of existing links (and adjusts the effort level). However, deviations where a pair of agents creates a link and/or adjusts their effort levels and *simultaneously* deletes any subset of existing links are not considered.

A strategy profile $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$ is a *strict pairwise Nash equilibrium (SPNE)* *iff*

- for any $i \in N$ and every $\mathbf{s}_i \in S_i$ such that $\bar{\mathbf{g}}(\mathbf{s}_i, \mathbf{s}_{-i}^*) \neq \bar{\mathbf{g}}^*(\mathbf{s}^*)$, $\Pi_i(\mathbf{s}^*) > \Pi_i(\mathbf{s}_i, \mathbf{s}_{-i}^*)$;
- for all $\bar{g}_{i,j} = 0$, if $\Pi_i(x'_i, x'_j, \mathbf{x}_{-i,-j}^*, \bar{\mathbf{g}}^* + \bar{g}_{i,j}) \geq \Pi_i(\mathbf{s}^*)$,
then $\Pi_j(x'_i, x'_j, \mathbf{x}_{-i,-j}^*, \bar{\mathbf{g}}^* + \bar{g}_{i,j}) < \Pi_j(\mathbf{s}^*)$.

Strict pairwise Nash equilibrium rules out unilateral and bilateral deviations that affect the network, but for which payoffs remain the same for the agents involved in a deviation.¹³

3 Analysis

3.1 Two-sided Network Formation

We start by presenting a result for the existence and uniqueness of a Nash equilibrium on a fixed network. All proofs are relegated to Appendix B. Note first that a necessary condition for $(\mathbf{x}, \bar{\mathbf{g}})$ to be a pairwise Nash equilibrium is that agents play Nash equilibrium effort levels on the network $\bar{\mathbf{g}}$.¹⁴ From strict strategic complementarities it follows directly that Nash equilibrium effort levels must be equal for all players in a complete component.¹⁵ Furthermore, Nash equilibrium effort levels strictly increase when adding a link to a connected network and effort levels are maximal in the complete network.

Proposition 1: *For any fixed network, $\bar{\mathbf{g}}$, there exists a unique NE in effort levels. Furthermore, (i) NE effort levels are equal for all players in a complete component, (ii) NE*

¹³For example, assume that \mathbf{g} is the profile of link announcements of a particular SPNE with undirected network $\bar{\mathbf{g}}$ and that $g_{i,j} = 1$ and $g_{j,i} = 0$. Assume that agent i deviates by changing the link announcements to j , so that $g_{i,j} = 0$. SPNE allows for this type of deviation, since $\bar{\mathbf{g}}$ is not altered and agent i 's payoff remains the same.

¹⁴For the case of linear best response functions we can apply the existence and uniqueness result provided by Ballester et al. (2006). For the case of concave best response functions we make use of a fixed point theorem provided by Kennan (2001). Bätz (2015) first applied Kennan's (2001) result in a network context.

¹⁵A component of a network is a subnetwork, such that any two nodes are connected to each other via paths, and no node in the subnetwork is connected to a node not in the subnetwork. A component is complete, if all nodes in the component are directly connected.

effort levels are maximal in the complete network and (iii) when adding a link to a connected component, NE effort levels of all agents in the component strictly increase.

At this point it is useful to emphasize what the (strict) convexity of the value function entails in our model: the incentives to link and access other players' effort levels are (strictly) increasing in the effort level already accessed. Note also that by invoking the value function we implicitly assume that deviating agents adjust their effort levels when creating a link.

We start by defining two cost threshold cost, $\underline{\kappa}$ and $\bar{\kappa}$. The lower threshold, $\underline{\kappa}$, is given by the gross marginal payoff when a pair of agents creates a link in the empty network, $\bar{\mathbf{g}}^e$. The higher threshold, $\bar{\kappa}$, is defined as the average gross marginal payoff of linking to $n - 1$ agents in the complete network, $\bar{\mathbf{g}}^c$. Both thresholds are expressed in terms of the value function. Denote the unique Nash equilibrium effort level in the complete network, $\bar{\mathbf{g}}^c$, by $x(\bar{\mathbf{g}}^c)$ and the unique Nash equilibrium effort level in the empty network, $\bar{\mathbf{g}}^e$, by $x(\bar{\mathbf{g}}^e)$. Furthermore, denote the corresponding vectors of Nash equilibrium effort levels with $\mathbf{x}(\bar{\mathbf{g}}^c)$ and $\mathbf{x}(\bar{\mathbf{g}}^e)$.

Definition 1: $\underline{\kappa} = v_i(x'_j(\bar{\mathbf{g}}^e + \bar{g}_{i,j})) - v(0)$ and $\bar{\kappa} = \frac{v((n-1)x(\bar{\mathbf{g}}^c)) - v(0)}{n-1}$.

As mentioned above, $\underline{\kappa}$ is given by the gross marginal value of two singleton agents creating a link. Note that, due to the convexity of the value function and strategic complementarities, this is also the lowest possible marginal value of creating a link. Therefore, for linking cost lower than $\underline{\kappa}$ the unique PNE is the complete network and for linking cost weakly higher than $\underline{\kappa}$ there exists a PNE such that the network is empty. Conversely, $\bar{\kappa}$ is defined as the average gross marginal payoff of sustaining $n - 1$ links in the complete network. Since effort levels are maximal in the complete network and due to the convexity of the value function, the highest linking cost can be sustained in the complete network. Moreover, if an agent in the complete network finds it profitable to delete any subset of links, then the agent also finds it profitable to delete all links (again due to the convexity of the value function). Therefore, for linking cost strictly higher than $\bar{\kappa}$ the unique PNE network is the empty network, while for linking cost weakly smaller than $\bar{\kappa}$, there exists a PNE such that the network is complete. The above is summarized in Proposition 2.

Proposition 2: $\underline{\kappa} < \bar{\kappa}$ holds. Furthermore, (i) if $\kappa < \underline{\kappa}$ then $(\mathbf{x}(\bar{\mathbf{g}}^c), \bar{\mathbf{g}}^c)$ is the unique PNE, (ii) if $\kappa > \bar{\kappa}$ then $(\mathbf{x}(\bar{\mathbf{g}}^e), \bar{\mathbf{g}}^e)$ is the unique PNE and (iii) if $\kappa \in [\underline{\kappa}, \bar{\kappa}]$ then $(\mathbf{x}(\bar{\mathbf{g}}^c), \bar{\mathbf{g}}^c)$ and $(\mathbf{x}(\bar{\mathbf{g}}^e), \bar{\mathbf{g}}^e)$ are PNE.

Next we formally define nested split graphs, which are a strict subset of core-periphery networks.^{16,17} Note that the star, the complete and the empty network are nested split graphs.

Definition 2: A network $\bar{\mathbf{g}}$ is a *nested split graph* if and only if

$$[\bar{g}_{i,l} = 1 \text{ and } \eta_k(\bar{\mathbf{g}}) \geq \eta_l(\bar{\mathbf{g}})] \Rightarrow \bar{g}_{i,k} = 1.$$

¹⁶A network $\bar{\mathbf{g}}$ is a core-periphery network if the set of agents N can be partitioned into two sets, $C(\bar{\mathbf{g}})$ (the core) and $P(\bar{\mathbf{g}})$ (the periphery), such that $\bar{g}_{i,j} = 1 \forall i, j \in C(\bar{\mathbf{g}})$ and $\bar{g}_{i,j} = 0 \forall i, j \in P(\bar{\mathbf{g}})$.

¹⁷For a formal proof that all nested split graphs are core-periphery networks see Chvátal and Hammer (1977).

Note that, in order to obtain pairwise Nash equilibrium networks such that agents' neighborhoods are not trivially nested, i.e. there exists a pair of agents, such that one agent's set of neighbors is a strict, non-empty subsets of another agent's set of neighbors, we need that the value function is strictly convex. That is, we need that agents' incentives to create links are strictly increasing in the effort level accessed. We first provide intuition for this case and then comment on the case when the value function is linear separately. Note first that, due to strategic complementarities, agents who access higher effort levels also exert higher effort levels. Due to the strict convexity of the value function, agents who access higher effort levels benefit more from linking to any particular agent. Conversely, agents prefer to link to agents with higher effort levels. Therefore, in any pairwise Nash equilibrium, agents with higher effort levels accessed (and therefore higher own effort levels) must be linked to all agents to which agents with lower effort levels accessed (and therefore lower own effort levels) are linked to. It is this reinforcing mechanism that can generate non-trivial nestedness. Agents with higher effort levels display a higher number of links for the same reason. As the network is nested in any pairwise Nash equilibrium, a higher number of links also implies a higher effort level accessed and, since the value function is increasing, higher gross payoffs.

If the value function is linear, then we can further refine the set of pairwise Nash equilibria. The network is then either complete, empty or a dominant group network.¹⁸ Note that in a dominant group network agents' neighborhoods are trivially nested in the sense that a pair of agents either shares the same neighbors (if both agents are in the complete component), or at least one agent is a singleton. The intuition for this result is that with a linear value function, the marginal value of linking to a particular agent is independent of the effort level already accessed. That is, if one agent finds it profitable to link to a particular agent, then all agents find it profitable. Since each agent in a connected component displays at least one link, there can then not exist a pair of agents that is not directly connected in a connected component. Therefore, any connected component must be complete.

Finally, we can refine the set of pairwise Nash equilibria slightly by showing that for any pair of agents the difference in the number of links must be different from one. Note that, although there are possibly many pairwise Nash equilibria, nested split graphs constitute a small fraction of all possible network structures, already for small n . Moreover, this fraction converges to zero rapidly.¹⁹

Theorem 1: *In any PNE, $(\mathbf{x}, \bar{\mathbf{g}})$, the network $\bar{\mathbf{g}}$ is a nested split graph such that*

- (i) $x_i < x_k \Leftrightarrow \eta_i(\bar{\mathbf{g}}) < \eta_k(\bar{\mathbf{g}}) \Leftrightarrow \pi_i < \pi_k$ and
- (ii) $|\eta_i(\bar{\mathbf{g}}) - \eta_j(\bar{\mathbf{g}})| \neq 1 \ \forall i, j \in N$.

Corollary 1: *If $v(y)$ is linear, then in any PNE, $(\mathbf{x}, \bar{\mathbf{g}})$, the network $\bar{\mathbf{g}}$ is either complete, empty or a dominant group network (which satisfies Theorem 1).*

¹⁸A network $\bar{\mathbf{g}}$ is a dominant group network if there exists one complete component with $1 < k < n$ agents, while $n - k > 0$ agents are isolated (Goyal, 2007).

¹⁹Hagberg et al. (2006) provide a algorithm to calculate the number of nested split graphs as a function of n , which can be written as $2^{(n-1)}$, while the number of all possible networks is given by $2^{\binom{n-1}{2}}$. That is, for $n = 6$ only 0.01% of all graphs are nested split graphs, while for $n = 7$ this number already drops to 0,003%.

Next we present an example to illustrate that with a strictly convex value function our model can generate pairwise Nash equilibria with multiple equilibrium effort levels and rich network structures. We depict two networks, $\bar{\mathbf{g}}_1$ and $\bar{\mathbf{g}}_2$, which can be sustained as *PNE* for the same parameter values. Note also that neither $\bar{\mathbf{g}}_1 \subset \bar{\mathbf{g}}_2$, nor $\bar{\mathbf{g}}_2 \subset \bar{\mathbf{g}}_1$ holds.

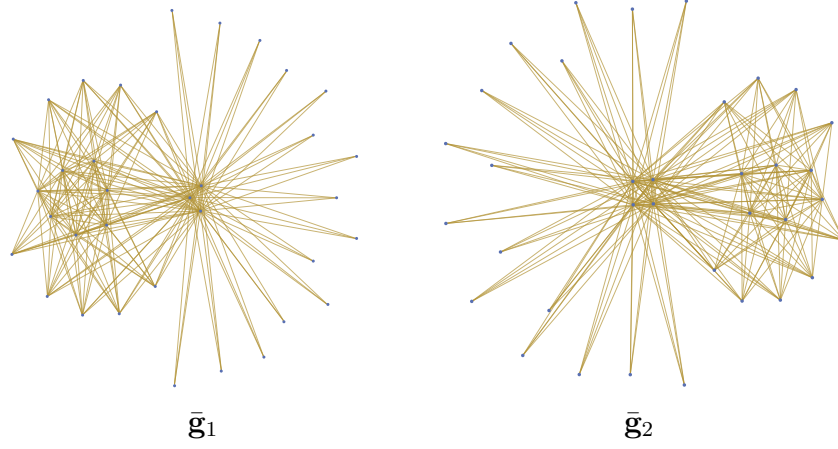


Figure 2: *PNE* networks $\bar{\mathbf{g}}_1$ and $\bar{\mathbf{g}}_2$ (Example 3)

Example 3: Assume that payoffs are given by $\pi_i(\mathbf{x}, \bar{\mathbf{g}}) = x_i - \frac{\beta}{2}x_i^2 + \lambda x_i \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$ and that $n = 35$, $\beta = \frac{1}{10}$, $\lambda = \frac{1}{1000}$ and $\kappa = \frac{13}{100}$. Then the networks $\bar{\mathbf{g}}_1$ and $\bar{\mathbf{g}}_2$ are *PNE* networks. Due to the strict convexity of the value function, the incentives to create links are strictly lower for agents with lower effort level accessed (i.e. for agents with a lower number of links in a nested split graph), while incentives to maintain links are strictly higher for agents with higher effort level accessed (i.e. for agents with a higher number of links in a nested split graph). This allows for the possibility of multi-layered, not trivially nested *PNE* networks, as the ones depicted in Figure 2 above.

Corollary 2 follows directly from Proposition 1 and Proposition 2 and points out similarities with supermodular games, such as the presence of a maximal and minimal pairwise Nash equilibria. Note, however, that our network formation game is not supermodular, neither when assuming that a link announcement needs to be reciprocated to form a link, nor in the one-sided specification of the model. Corollary 2 also emphasizes that, if there exists a pairwise Nash equilibrium that is neither empty nor complete, then both, the empty and the complete network, can be sustained as pairwise Nash equilibria. Note further that individual payoffs are shown to be maximal in the pairwise Nash equilibrium such that the network is complete. In our setting static coordination failures may prevent agents from reaching this structure. Note that this is also true for the one-sided model, where deviating agents may create multiple links.²⁰

Corollary 2: *If there exists a PNE, $(\mathbf{x}, \bar{\mathbf{g}})$, such that $\bar{\mathbf{g}} \notin \{\bar{\mathbf{g}}^c, \bar{\mathbf{g}}^e\}$, then (i) $(\mathbf{x}(\bar{\mathbf{g}}^e), \bar{\mathbf{g}}^e)$ and $(\mathbf{x}(\bar{\mathbf{g}}^c), \bar{\mathbf{g}}^c)$ are also PNE, (ii) $(\mathbf{x}(\bar{\mathbf{g}}^e), \bar{\mathbf{g}}^e)$ is the minimal and $(\mathbf{x}(\bar{\mathbf{g}}^c), \bar{\mathbf{g}}^c)$ is the maximal PNE, (iii) $(\mathbf{x}(\bar{\mathbf{g}}^c), \bar{\mathbf{g}}^c)$ is the PNE where individual payoffs are maximal.*

²⁰Dutta et al. (2005) show in a dynamic model without effort choice and under conditions that are reminiscent of a convex value function the following interesting result: Even when agents are farsighted the uniquely efficient structure (the complete network) may not be reached in equilibrium.

We provide necessary and sufficient conditions for a strict pairwise Nash equilibrium star network for the linear-quadratic payoff specification.²¹ The star network is of interest, as it relates our results to Galeotti and Goyal (2010), who obtain star structures in the presence of strategic substitutes. Next, recall from Theorem 1 that in any pairwise Nash equilibrium more central agents, which in a nested split graph coincide with agents with a higher number of links, obtain higher gross payoffs. Below we provide an example, which illustrates that a more central agent may, however, obtain strictly lower net payoffs. In this case the cost of sustaining a higher number of links outweighs higher gross payoffs.

Example 4: Assume the linear-quadratic gross payoff function $\pi_i(\mathbf{x}, \bar{\mathbf{g}}) = x_i - \frac{1}{2}\beta x_i^2 + \lambda x_i \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$ and $n = 10$, $\beta = 1$, $\lambda = \frac{1}{10}$ and $\kappa = \frac{18}{100}$. Then the star is a pairwise Nash equilibrium and net payoffs for the central agent are approximately 0.38, while for agents in the periphery they are approximately 0.55.

3.2 Multiplicity of pairwise Nash equilibria

Next we shed light on equilibrium multiplicity. Recall that if $\bar{\mathbf{g}}' \subset \bar{\mathbf{g}}$ holds, then all links that are present in $\bar{\mathbf{g}}'$ are also present in $\bar{\mathbf{g}}$ and there are some links that are present in $\bar{\mathbf{g}}$ but not in $\bar{\mathbf{g}}'$. For the following assume $\bar{\mathbf{g}}$ to be a nested split graph that satisfies the requirements of Theorem 1. Proposition 3 then provides conditions such that, if there exists a pair of pairwise Nash equilibria with networks $\bar{\mathbf{g}}'$ and $\bar{\mathbf{g}}''$, such that neither $\bar{\mathbf{g}}'$ nor $\bar{\mathbf{g}}''$ are empty or complete and $\bar{\mathbf{g}}' \subset \bar{\mathbf{g}} \subset \bar{\mathbf{g}}''$ holds, then $\bar{\mathbf{g}}$ is also a pairwise Nash equilibrium network. That is, conditions are provided such that a network $\bar{\mathbf{g}}$, which may be considered an intermediate network relative to the *PNE* networks $\bar{\mathbf{g}}'$ and $\bar{\mathbf{g}}''$, is also a *PNE* network. We illustrate Proposition 3 by way of a simple example. Call a complete core-periphery network a core-periphery network such that all links between the core and the periphery are in place. For example, a complete core-periphery network with a core of size 1 is a star network such that the center displays $n - 1$ links. Assume that there are ten agents and two pairwise Nash equilibria, such that one displays a complete core-periphery network with a core of size three and another with a core of size six. From Proposition 3 it then follows that there also exist pairwise Nash equilibria with a complete core-periphery network such that the core is of size four and such that the core is of size five.²²

To show our more general result we first define a preorder, denoted by \succsim , which allows us to compare certain links across networks. We then use the preorder \succsim to infer that if certain deviations are not profitable in $\bar{\mathbf{g}}'$, then they are also not profitable in $\bar{\mathbf{g}}$ and, likewise, if certain deviations are not profitable in $\bar{\mathbf{g}}''$, then they are also not profitable in $\bar{\mathbf{g}}$.

Definition 3: $\bar{g}_{i,j} \succsim \bar{g}'_{k,l}$ iff $\max\{\eta_i(\bar{\mathbf{g}}), \eta_j(\bar{\mathbf{g}})\} \geq \max\{\eta_k(\bar{\mathbf{g}}'), \eta_l(\bar{\mathbf{g}}')\}$ and $\min\{\eta_i(\bar{\mathbf{g}}), \eta_j(\bar{\mathbf{g}})\} \geq \min\{\eta_k(\bar{\mathbf{g}}'), \eta_l(\bar{\mathbf{g}}')\}$ holds.

²¹These are given by $\beta > \lambda(2 + \sqrt{2})$, $1 + \frac{\beta^2(5\beta-3\lambda)}{(\beta-\lambda)^3} < n < 1 + \frac{\beta^2}{\lambda^2}$ and $\frac{\beta(2\beta-\lambda)\lambda(\beta+\lambda)^2}{2(\beta-\lambda)^2(\beta^2-(n-1)\lambda^2)^2} < \kappa < \frac{\lambda(2\beta^3+(n+1)\beta^2\lambda-(n-1)\lambda^3)}{2\beta(\beta^2-(n-1)\lambda^2)^2}$. The conditions were derived in Mathematica and are available from the author upon request.

²²A simple algorithm to calculate the set of all pairwise Nash equilibrium networks for the linear-quadratic payoff specification (and for given parameter values) is available from the author upon request. The results from the equilibrium characterization, together with properties of nested split graphs, are used to reduce the number of deviations that need to be checked for a nested split graph to also be a pairwise Nash equilibrium network.

Proposition 3: *If there exist a pair of PNE, $(\mathbf{x}', \bar{\mathbf{g}}')$ and $(\mathbf{x}'', \bar{\mathbf{g}}'')$, such that $\bar{\mathbf{g}}' \subset \bar{\mathbf{g}}''$ and $\bar{\mathbf{g}}', \bar{\mathbf{g}}'' \notin \{\bar{\mathbf{g}}^e, \bar{\mathbf{g}}^c\}$, then any $\bar{\mathbf{g}}$ with $\bar{\mathbf{g}}' \subset \bar{\mathbf{g}} \subset \bar{\mathbf{g}}''$ is also a PNE network if the following conditions (and the ones given in Theorem 1) hold*

- i) $\forall i, j$ such that $\bar{g}_{i,j} = 1$ and $\bar{g}'_{i,j} = 0$, there exists a k, l such that $\bar{g}'_{k,l} = 1$ with $\bar{g}_{i,j} \succ \bar{g}'_{k,l}$,
- ii) $\forall i, j$ such that $\bar{g}_{i,j} = 0$ and $\bar{g}''_{i,j} = 1$, there exists a k, l such that $\bar{g}''_{k,l} = 0$ with $\bar{g}''_{k,l} \succ \bar{g}_{i,j}$.

3.3 Robustness to Heterogeneity

In this section we allow for heterogeneity in gross payoff functions and individual linking cost. More specifically, assume that each agent's gross payoff function is parametrized by a vector $\boldsymbol{\theta}_i \in \Theta$ with $\Theta \subseteq \mathbb{R}^k$ and Θ open, so that an agent i 's gross payoff function is given by $\pi_i(x, y, \boldsymbol{\theta}_i) = \pi(x, y, \boldsymbol{\theta}_i)$. We now allow for $\boldsymbol{\theta}_i \neq \boldsymbol{\theta}_j$ and define $\boldsymbol{\theta}$ by $\boldsymbol{\theta} = \boldsymbol{\theta}_1 \times \dots \times \boldsymbol{\theta}_n$. Assume further that the assumptions on payoffs from the homogeneous agent case are satisfied for all $\boldsymbol{\theta}_i \in \Theta$. Moreover, we assume that payoff functions are such that for all $\boldsymbol{\theta}_i \in \Theta$, best response functions are either all strictly concave or all linear.²³ When considering network formation, we also allow for differences in individual linking cost, i.e. $\kappa_i \neq \kappa_j$, and denote the vector of individual linking cost by $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_n)$.

We start by showing in Proposition 4 that, if the conditions for existence and uniqueness of a Nash equilibrium in effort levels on a fixed network for homogeneous agents hold individually for each agent, then existence and uniqueness is also guaranteed in the heterogeneous agent case. With the exception of (i), Proposition 1 extends to heterogeneous agents.

Proposition 4: *If agents are heterogeneous, then for any fixed network, $\bar{\mathbf{g}}$, there exists a unique NE in effort levels. Furthermore, NE effort levels are maximal in the complete network and, when adding a link to a connected component, NE effort levels of all agents in the component strictly increase.*

We next show that if a pairwise Nash equilibrium $(\mathbf{x}, \bar{\mathbf{g}})$ is strict in the model with homogeneous agents and sufficiently small heterogeneity is introduced, then generically there exists a (strict) pairwise Nash equilibrium $(\tilde{\mathbf{x}}, \bar{\mathbf{g}})$, such that the vector of equilibrium effort levels may be altered, but the network remains unchanged. That is, strict pairwise Nash equilibrium networks are generically robust to the introduction of sufficiently small heterogeneity. To show our result, we need to further assume that the payoff function for the homogeneous agent case, π , is (jointly) continuous and smooth. Note that these conditions are satisfied for the linear-quadratic model specification. Note further that conditions for the existence of a SPNE star network are provided for the linear-quadratic case in Section 3.1. The proof first uses the Theorem of the Maximum to show that each agents' best response and value functions are continuous in the vector of parameters $\boldsymbol{\theta}_i$ (and effort level accessed y_i). We then employ the Implicit Function Theorem and Sard's Theorem to show that, generically, Nash equilibrium effort levels on a network $\bar{\mathbf{g}}$ also change continuously in $\boldsymbol{\theta}$. Since best response functions and value functions change continuously in $\boldsymbol{\theta}$, gross payoffs of a deviation also change continuously in $\boldsymbol{\theta}$. Finally, note that equilibrium and deviation payoffs change continuously in $\boldsymbol{\kappa}$. That is, equilibrium and deviation payoffs change continuously in $\boldsymbol{\theta}$ and $\boldsymbol{\kappa}$. Therefore, strict pairwise Nash equilibrium networks are generically

²³We can extend the derivations for the conditions on payoffs from Appendix A. That is, we assume that either $-\partial(\frac{\partial^2 \pi(x, y, \boldsymbol{\theta}_i)}{\partial x \partial y} / \frac{\partial^2 \pi(x, y, \boldsymbol{\theta}_i)}{\partial^2 x}) / \partial y \big|_{x=\bar{x}(y, \boldsymbol{\theta}_i)} < 0$ holds $\forall y$ and $\forall \boldsymbol{\theta}_i \in \Theta$, or $-\partial(\frac{\partial^2 \pi(x, y, \boldsymbol{\theta}_i)}{\partial x \partial y} / \frac{\partial^2 \pi(x, y, \boldsymbol{\theta}_i)}{\partial^2 x}) / \partial y_i \big|_{x=\bar{x}(y, \boldsymbol{\theta}_i)} = 0$ holds $\forall y$ and $\forall \boldsymbol{\theta}_i \in \Theta$.

robust to the introduction of sufficiently small heterogeneity in gross payoff functions and individual linking cost. For convenience, formal definitions of the Theorem of the Maximum, the Implicit Function Theorem and Sard's Theorem are provided in Appendix B.

Proposition 5: *A SPNE network is generically robust to the introduction of sufficiently small heterogeneity in gross payoffs and individual linking cost.*

However, if heterogeneity in best response functions is sufficiently large, then pairwise Nash equilibrium networks may be different from nested split graphs. We present an example with two types and a *PNE* network that consists of two separate components and is therefore not a nested split graph. Similar examples can easily be constructed when allowing for heterogeneous linking cost.

Example 5: Assume that there are two type of agents with corresponding gross payoff functions $\pi_i(\mathbf{x}, \bar{\mathbf{g}}, \boldsymbol{\theta}_i) = x_i - \frac{1}{2}\beta_i x_i^2 + \lambda_i x_i \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$, where $\boldsymbol{\theta}_i = (\beta_i, \lambda_i)$ and $\boldsymbol{\theta}_i \in \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2\}$ such that $\boldsymbol{\theta}_1 = (1, \frac{1}{100})$, while $\boldsymbol{\theta}_2 = (2, \frac{1}{10})$ and $\kappa = \frac{1}{100}$. Assume there are 15 agents of type $\boldsymbol{\theta}_1$ and 3 agents of type $\boldsymbol{\theta}_2$. The network where agents of type $\boldsymbol{\theta}_1$ form a complete component and agents of type $\boldsymbol{\theta}_2$ form another complete component with no links across components is a *PNE* network. Note that in the *PNE* considered, type $\boldsymbol{\theta}_2$ agents exert less effort than type $\boldsymbol{\theta}_1$ agents. Furthermore, while type $\boldsymbol{\theta}_2$ agents find it profitable to link to each other and would profit from linking to type $\boldsymbol{\theta}_1$ agents, this deviation is not profitable for type $\boldsymbol{\theta}_1$ agents.

3.4 Efficiency

Next, we turn to social welfare. We assume that the total cost of the network (*the network cost*) is given by $\Phi(\eta(\bar{\mathbf{g}}))$, where $\eta(\bar{\mathbf{g}})$ is the total number of links in network $\bar{\mathbf{g}}$. Assume that $\Phi(0) = 0$. To simplify notation we sometimes write $\Phi(\bar{\mathbf{g}})$ for $\Phi(\eta(\bar{\mathbf{g}}))$ and η for $\eta(\bar{\mathbf{g}})$. The case of linear linking cost κ corresponds to Φ being linear and $\Phi(\eta) - \Phi(\eta - 1) = 2\kappa \forall \eta$.

Social welfare is defined as the sum of individual gross payoffs minus the network cost. For any strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ and resulting network, $\bar{\mathbf{g}}$, social welfare is given by

$$W(\mathbf{x}, \mathbf{g}) = \sum_{i \in N} \pi_i(\mathbf{x}, \bar{\mathbf{g}}) - \Phi(\bar{\mathbf{g}}).$$

A strategy profile $\hat{\mathbf{s}}$ is socially efficient if $W(\hat{\mathbf{s}}) \geq W(\mathbf{s}) \forall \mathbf{s} \in S$. Denote the vector of efficient effort levels for a given network $\bar{\mathbf{g}}$ with $\hat{\mathbf{x}}(\bar{\mathbf{g}})$ and agent i 's entry in vector $\hat{\mathbf{x}}(\bar{\mathbf{g}})$ by $\hat{x}_i(\bar{\mathbf{g}})$. That is, $\hat{\mathbf{x}}(\bar{\mathbf{g}})$ yields the highest sum of payoffs for a given network, $\bar{\mathbf{g}}$, so that $W(\hat{\mathbf{x}}(\bar{\mathbf{g}}), \bar{\mathbf{g}}) \geq W(\mathbf{x}(\bar{\mathbf{g}}), \bar{\mathbf{g}}) \forall \mathbf{x} \in \mathbb{R}_+^n$. To simplify notation, we sometimes write $\hat{\mathbf{x}}$ for $\hat{\mathbf{x}}(\bar{\mathbf{g}})$ and \hat{x}_i for $\hat{x}_i(\bar{\mathbf{g}})$. Denote the efficient (undirected) network by $\hat{\mathbf{g}}$, so that $W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}}) \geq W(\hat{\mathbf{x}}(\bar{\mathbf{g}}), \bar{\mathbf{g}}) \forall \bar{\mathbf{g}} \in G$. To specify a well defined problem, we assume that optimal effort levels and payoffs are bounded in the complete network and therefore bounded in any other network. A sufficient condition is that $\pi(x, (n-1)x)$ is strictly concave in x and $\partial \pi(x, (n-1)x) / \partial x = 0$ for some x .²⁴

In Theorem 2 we show that if $\partial^2 \pi(x, y) / \partial^2 y \geq 0$ holds $\forall x, y$, then every efficient network is a nested split graph.²⁵ This again subsumes the linear-quadratic setup but also allows for

²⁴For the linear-quadratic payoff specification these conditions are satisfied when $2\lambda / (2\lambda + \beta) < 1/n$.

²⁵The case when $\partial^2 \pi(x, y) / \partial^2 y < 0$ remains an open question. We ran extensive numerical simulations and, if the value function is convex, then for all parameter values considered all efficient networks are nested split graphs. If $\partial^2 \pi(x, y) / \partial^2 y < 0$ and the value function is concave, then one can easily find examples such that the efficient networks are not nested split graphs (e.g. the circle).

concave best response functions. Note that when $\partial^2\pi(x,y)/\partial^2y \geq 0$ holds $\forall x,y$, then the value function is strictly convex. To show our result we consider rewiring links (similar to Belhaj et al., 2016) and simultaneous adjustments of effort levels. Multiple links may need to be rewired to ensure that, not only are agents with higher effort levels connected, but also that agents with higher own effort levels access more effort. This maximizes positive externalities. We thereby obtain an alignment in terms of effort levels, effort levels accessed and number of links, which mirrors our equilibrium characterization and yields nested split graphs as the only efficient networks. Note, however, that with the exception of the empty network and high linking cost, pairwise Nash equilibria are typically not efficient. Not only are, due to positive externalities, effort levels suboptimally low, but efficient networks may also differ from *PNE* networks.²⁶

Theorem 2: *If $\frac{\partial^2\pi(x,y)}{\partial^2y} \geq 0$ holds $\forall x,y$, then $\hat{\mathbf{g}}$ is a nested split graph and*

- (i) $\hat{x}_i(\hat{\mathbf{g}}) > \hat{x}_j(\hat{\mathbf{g}}) \Leftrightarrow \hat{y}_i(\hat{\mathbf{g}}) > \hat{y}_j(\hat{\mathbf{g}}) \Leftrightarrow \hat{\eta}_i(\hat{\mathbf{g}}) > \hat{\eta}_j(\hat{\mathbf{g}})$,
- (ii) *if Φ is weakly concave, then the efficient network is either empty or complete,*
- (iii) *if Φ is strictly convex, $\Phi(1) \simeq 0$ and $\Phi(\frac{n(n-1)}{2}) > \sum_{i \in N} \pi_i(\hat{\mathbf{x}}(\bar{\mathbf{g}}^c), \bar{\mathbf{g}}^c)$, then the efficient network is a nested split graph that is neither empty nor complete.*

We present a simple example for an efficient configuration with 4 agents and a strictly convex network cost function.

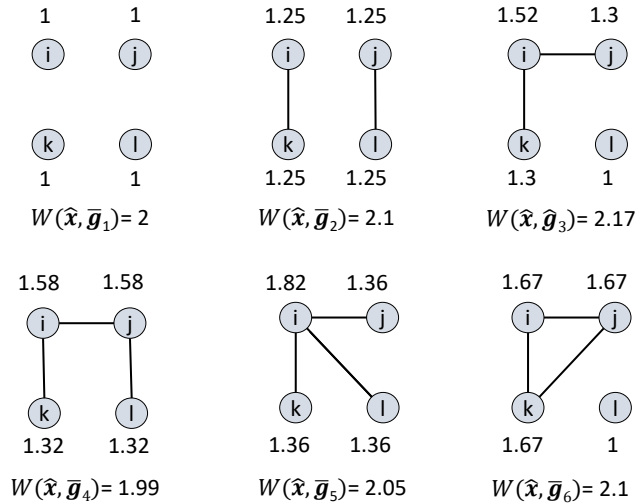


Figure 3: Optimal effort levels and total payoffs for different networks

Example 6: Assume $n = 4$ and payoffs are given by $\pi_i(\mathbf{x}, \bar{\mathbf{g}}) = x_i - \frac{1}{2}\beta x_i^2 + \lambda x_i \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$ with network cost function $\Phi(\bar{\mathbf{g}}) = c\eta(\bar{\mathbf{g}})^2$, where $\beta = 1$, $\lambda = 0.1$ and $c = 0.1$. We depict the relevant network structures above, together with efficient effort levels, $\hat{\mathbf{x}}$, given a particular network $\bar{\mathbf{g}}$ and the total sum of payoffs, $W(\hat{\mathbf{x}}(\bar{\mathbf{g}}), \bar{\mathbf{g}})$. Note that any network with four or

²⁶This will be the case, for example, when Φ is linear and $\kappa = \bar{\kappa} + \epsilon$ with ϵ small. Then, the unique *PNE* is such that the network is empty, while the efficient configuration is such that the network is complete. Note also that from Proposition 2 it follows directly that with Φ linear the second-best efficient network (i.e. the planner chooses the network, while agents play *NE* effort levels given the network) is complete for $\kappa < \bar{\kappa}$ and empty for $\kappa > \bar{\kappa}$.

more links yields payoffs smaller than in the empty network and we omit these networks in Figure 3. A star consisting of three agents and one singleton agent, network $\hat{\mathbf{g}}_3$, is the uniquely efficient network structure. Rewiring the link between j and l in $\bar{\mathbf{g}}_2$, which is not a nested split graph, to i and j and adjusting effort levels yields the efficient configuration $\hat{\mathbf{x}}(\hat{\mathbf{g}}_3)$, $\hat{\mathbf{g}}_3$. Note also that rewiring the network $\bar{\mathbf{g}}_4$, so that we obtain nested split graphs $\bar{\mathbf{g}}_5$ and $\bar{\mathbf{g}}_6$ and adjusting effort levels, strictly increases total payoffs.

4 Discussion

In this section we briefly comment on the case of one-sided network formation, which is presented in the online appendix. Furthermore, a simple example with a concave value function and two-sided network formation is provided, such that a pairwise Nash equilibrium network exists that does not display a core-periphery structure and is therefore not a nested split graph.

To show that our results are not driven by whether link formation is one-sided or two-sided, we characterize Nash equilibria for the one-sided network formation case. More precisely, if the value function is strictly convex, then again all equilibrium networks are nested split graphs. In the following we describe the model informally. Note first that under one-sided network formation agents may unilaterally create and delete multiple links, potentially simultaneously, and adjust their effort levels. The agent extending the link incurs the cost of the link, while both agents involved in a link access the effort level of the respective other agent. However, in the deviations considered, agents receiving a link do not adjust their effort level. Essentially, this is why we need to assume a strictly convex value function to show our result in the one-sided case. The mechanism is similar to the two-sided network formation case. Not only does linking to agents with higher effort levels (and therefore higher effort level accessed) yield higher marginal payoffs, but agents with higher effort levels also find it more profitable to link to any given agent. Again this induces nested network structures in equilibrium.

The characterization of pairwise Nash equilibria with two-sided network formation and a concave value function is outside the scope of this paper. Below we provide an example with a concave value function, such that the pairwise Nash equilibrium network is neither a nested split graph, nor a core-periphery network.

Example 7: Assume payoffs are given by $\Pi_i(\mathbf{s}) = 2\sqrt{x_i \sqrt{(\sum_{j \in N_i(\bar{\mathbf{g}})} x_j)} - x_i - \eta_i(\bar{\mathbf{g}})\kappa}$, which is a special case of the baseline model in Bätz (2015). Assume $n = 4$. The circle is a pairwise Nash equilibrium for $\kappa \in [\frac{1}{2}(\sqrt{17} - 3), 2 - \sqrt{2}]$.

5 Contribution and Relation to the Networks Literature

In the following we comment on our contribution and how the paper relates to the literature. Building on the seminal contribution by Ballester et al. (2006), a theoretical and empirical body of work studies effort choice under strategic complementarities and positive externalities on fixed networks. Payoffs are assumed to be linear-quadratic, while best response functions are linear. See, for example, Calvó-Armengol et al. (2005 and 2009). The linear-quadratic payoff specification implicitly assumes a convex value function and we therefore endogenize

the network for the homogeneous agents case of this prominent strand of the literature. While the focus of our paper is on network formation, we also provide results for the existence and uniqueness of equilibrium on fixed networks for linear as well as concave best response functions. Important work on this subject matter includes Allouch (2015), Belhaj et al. (2014) and Bramoullé et al. (2014).

Our paper contributes to a literature, which studies endogenous network formation with effort choice in one-shot, simultaneous move games. Galeotti and Goyal (2010) present a model with strategic substitutes and one-sided network formation.²⁷ We show that core-periphery networks (and in particular star networks), as obtained in Galeotti and Goyal (2010), are not a feature of strategic substitutes alone but may also be sustained with strategic complements. Bätz (2015) assumes strategic complementarities and a concave value function. The author shows that under one-sided network formation a wide range of different network structures may be sustained in equilibrium, such as the empty, complete and regular networks, biregular bipartite graphs, multipartite graphs, and, under fairly restrictive assumptions on payoffs, also core-periphery networks. One of our main contributions is to show, for the one-sided and the more involved two-sided network formation case, that, if the value function is convex, then all equilibria are nested split graphs.²⁸

A separate literature studies network formation and action choices in a dynamic setting with myopic agents. In a recent paper König et al. (2014) employ the linear-quadratic specification. At each time period agents are assumed to play a two-stage game. In the first stage agents play equilibrium effort levels on a fixed network. In the second stage one agent is chosen at random, who myopically updates the linking strategy in one dimension, i.e. by creating a single, one-sided link at zero cost. Links are assumed to decay over time, with more valuable links decaying at a slower rate. The authors show that, if the rate at which agents may create a one-sided link is intermediate, then all networks obtained under perturbed and unperturbed dynamics are nested split graphs (which are neither empty nor complete).²⁹ König et al. (2014) then test their model empirically for banking and trade networks and find evidence for nestedness. Note that the payoff function in König et al. (2014) is a simplified version of the payoff function considered in Ballester et al. (2006) and equilibrium effort levels are again directly proportional to Bonacich centrality. In fact all results in König et al. (2014), as the authors point out, remain unchanged if agents were to form links according to measures of centrality other than Bonacich centrality, such as degree, closeness, or betweenness centrality. In contrast, we show in this paper that in the presence of linking cost (and simultaneous moves) nested split graphs can be sustained as (pairwise) Nash equilibria of a link formation game. One interpretation is that our paper provides game-theoretic support for nested split graphs, as obtained in König et al. (2014). Furthermore, we show that the relevant condition for all equilibrium structures to be nested split graphs is given by a convex value function.

Belhaj et al. (2016) are concerned with efficiency. The authors study the problem of

²⁷For a one-sided network formation model with strategic substitutes and heterogeneous agents, see Kinader and Merlino (forthcoming).

²⁸Cabrales et al. (2011) present a related model with a convex value function, but with a very different approach. The authors conceive network formation as an untargeted socialization effort.

²⁹Lagerås and Seim (2016) present a model that has similar features as König et al. (2014) and again stochastically stable networks are nested split graphs. Note that for link creation Lagerås and Seim (2016) require consent of both agents. However, since the authors assume zero linking cost and impose strict positive externalities, an agent always accepts a link that is to be created. This amounts to a model of one-sided network formation.

a planner, who chooses the network and agents play Nash equilibrium effort levels, taking the network as given. Considering linear-quadratic payoffs and allowing for general network cost functions, Belhaj et al. (2016) find that all efficient networks are nested split graphs. The authors then show that in the linear-quadratic setting their results carry over to the case when the planner may also choose effort levels. Another contribution of our paper is to demonstrate that the curvature of the value function also bears relevance for efficiency considerations. More specifically, we show that under additional assumptions on payoffs, which allow for concave best-response functions, again all first-best efficient networks are nested split graphs. That is, we generalize some of the results of Belhaj et al. (2016) and, different from Belhaj et al. (2016), our proof directly considers the problem of a planner who chooses both the network and effort levels.

6 Conclusion

Peer effects and social structure are important for the understanding of many social and economic phenomena, such as educational attainment, criminal activity, labor market participation, and R&D expenditures of firms. We present a simple model of strategic network formation in the presence of peer effects, which sheds light on the emergence of frequently observed patterns in terms of network structure, effort and payoffs/performance. The model is sufficiently rich to be descriptive of many different networks, which have been observed in a variety of settings, while at the same time capturing a common structural feature: nested neighborhoods. A large class of payoff functions is considered, which includes the homogeneous agent case of the linear-quadratic specification, frequently used in theoretical and empirical work. Different from much of the prevailing literature, we allow for concave best-response functions and consider two-sided network formation, while also characterizing the one-sided case. We provide the counter-intuitive result that more central agents may obtain strictly lower net payoffs (net of linking cost). This has potential implications for empirical work, which typically disregards linking cost. One of the theoretical insights drawn from our analysis is that, if the value function is convex, then *all* (pairwise) Nash equilibria are nested split graphs. Furthermore, the curvature of the value function is also shown to be important for efficiency considerations.

7 Appendix A

The conditions on the payoff function $\pi(x, y)$ are as follows. We assume $\frac{\partial^2 \pi(x, y)}{\partial^2 x} < 0$, which (together with the convexity of X) guarantees a unique maximizer, denoted by $\bar{x}(y)$. Strict positive local externalities and strict strategic complementarities are given by $\frac{\partial \pi(x, y)}{\partial y} > 0$ and $\frac{\partial^2 \pi(x, y)}{\partial x \partial y} > 0 \forall x, y$. Note that the necessary condition for a maximum is given by

$$\frac{\partial \pi(x, y)}{\partial x} \big|_{x=\bar{x}(y)} = 0.$$

By the implicit function theorem we obtain

$$\frac{d\bar{x}(y)}{dy} = -\left(\frac{\partial^2 \pi(x, y)}{\partial x \partial y} / \frac{\partial^2 \pi(x, y)}{\partial^2 x}\right) \big|_{x=\bar{x}(y)}.$$

As \bar{x} is a maximum $\frac{\partial^2 \pi(x, y)}{\partial^2 x} \big|_{x=\bar{x}(y)} < 0$ holds. For $\frac{d\bar{x}(y)}{dy} > 0$ to hold we therefore need that $\frac{\partial^2 \pi(x, y)}{\partial x \partial y} \big|_{x=\bar{x}(y)} > 0 \forall y$. For the best response function to be concave we then need

$$-\partial\left(\frac{\partial^2 \pi(x, y)}{\partial x \partial y} / \frac{\partial^2 \pi(x, y)}{\partial^2 x}\right) / \partial y \big|_{x=\bar{x}(y)} < 0 \forall y,$$

while for the best response function to be linear we need

$$-\partial\left(\frac{\partial^2 \pi(x, y)}{\partial x \partial y} / \frac{\partial^2 \pi(x, y)}{\partial^2 x}\right) / \partial y \big|_{x=\bar{x}(y)} = 0 \forall y.$$

8 Appendix B

As in Kennan's paper, we define a vector \mathbf{b} to be larger than a vector \mathbf{a} if and only if $b_i > a_i \forall i \in N$.

Proof of Proposition 1. For the first part of the statement we discern two cases. First we consider payoff functions such that best response functions are linear, which allow us to use the existence result provided by Ballester et al. (2006). Without loss of generality, we can write the best response function as $\bar{x}_i(\sum_{j \in N_i(\bar{\mathbf{g}})} x_j) = \frac{1}{\beta} + \frac{\lambda}{\beta} \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$. A NE exists and is unique if and only if $\beta < \lambda \frac{1}{\mu_1(\bar{\mathbf{g}})}$, where $\mu_1(\bar{\mathbf{g}})$ is the largest eigenvalue of the adjacency matrix of $\bar{\mathbf{g}}$. The largest eigenvalue for a graph lies between the following bounds $\max\{d_{avg}(\bar{\mathbf{g}}), \sqrt{d_{max}(\bar{\mathbf{g}})}\} \leq \mu_1(\bar{\mathbf{g}}) \leq d_{max}(\bar{\mathbf{g}})$, where $d_{max}(\bar{\mathbf{g}})$ is the maximum degree and $d_{avg}(\bar{\mathbf{g}})$ the average degree in network $\bar{\mathbf{g}}$.³⁰ The largest eigenvalue for a graph is then maximal and equal to $n - 1$ in the complete network, $\bar{\mathbf{g}}^c$. To guarantee existence of a unique NE we therefore only need $\frac{\lambda}{\beta} < \frac{1}{n-1}$ to hold. Next we consider payoff functions such that best response functions are strictly concave. Define the function $f_{\bar{\mathbf{g}}} : X^n \rightarrow X^n$ as

$$f_{\bar{\mathbf{g}}}(\mathbf{x}) = \begin{pmatrix} \bar{x}(\sum_{j \in N_1(\bar{\mathbf{g}})} x_j) \\ \vdots \\ \bar{x}(\sum_{j \in N_n(\bar{\mathbf{g}})} x_j) \end{pmatrix}.$$

³⁰See, for example, L. Lovasz, Geometric Representations of Graphs (2009).

From strategic complementarities we know that $\bar{x}(y)$ is strictly increasing and, together with strict concavity of $\bar{x}(y)$, f is increasing and strictly concave. We now apply Kennan's result (Theorem 3.3 in Kennan, 2001), which is restated here. Suppose f is an increasing and strictly concave function from \mathbb{R}^n to \mathbb{R}^n such that $f(\mathbf{0}) \geq \mathbf{0}$, $f(\mathbf{a}) > \mathbf{a}$ for some positive vector \mathbf{a} , and $f(\mathbf{b}) < \mathbf{b}$ for some vector $\mathbf{b} > \mathbf{a}$. Then f has a unique positive fixed point. Recall that we assume $\bar{x}(0) > 0$ and therefore $f(\mathbf{0}) > \mathbf{0}$. To see that there exists a vector \mathbf{a} such that $f(\mathbf{a}) > \mathbf{a}$, choose $\mathbf{a} = (\varepsilon_1, \dots, \varepsilon_n)$ such that $\varepsilon_i = \varepsilon < \frac{\bar{x}(0)}{n-1} \forall i \in N$ with $\varepsilon > 0$. The effort level accessed of an agent i with $\eta_i(\bar{\mathbf{g}})$ neighbors is given by $\eta_i(\bar{\mathbf{g}})\varepsilon$. Note that we then have $\bar{x}(\eta_i(\bar{\mathbf{g}})\varepsilon) > \bar{x}(0) > \eta_i(\bar{\mathbf{g}})\varepsilon$. The first inequality holds because $\bar{x}(y)$ is strictly increasing, while the second inequality follows from $\bar{x}(0) > \varepsilon(n-1)$ and $n-1 \geq \eta_i(\bar{\mathbf{g}})$. Next we show that there exists a vector \mathbf{b} with $f(\mathbf{b}) < \mathbf{b}$. Choose again a vector such that all entries are equal, i.e. $\mathbf{b} = (b_1, \dots, b_n)$ with $b = b_i \forall i \in N$. The condition $f(\mathbf{b}) < \mathbf{b}$ can be written as $\bar{x}(\eta_i(\bar{\mathbf{g}})b) < b \forall i \in N$. For b sufficiently large $\bar{x}(\eta_i(\bar{\mathbf{g}})b) < b \forall i \in N$ holds due to $\frac{\partial \bar{x}(y)}{\partial y} < \frac{1}{n-1}$ for some value of y , the strict concavity of $\bar{x}(y)$ and $n \geq 3$. To show that $\mathbf{b} > \mathbf{a}$, note that we can choose ε (and therefore \mathbf{a}) arbitrarily close to zero for $\bar{x}(\eta_i(\bar{\mathbf{g}})\varepsilon) > \bar{x}(0) > \eta_i(\bar{\mathbf{g}})\varepsilon$ to hold.

Part (i): Assume to the contrary that there exists a NE , \mathbf{x} , with a pair of players k and l in a complete component such that $x_k \neq x_l$ and, without loss of generality, that $x_k > x_l$. Note that in a complete component $N_k(\bar{\mathbf{g}}) \setminus \{l\} = N_l(\bar{\mathbf{g}}) \setminus \{k\}$ holds and therefore $x_k > x_l$ implies $y_l = \sum_{j \in N_l(\bar{\mathbf{g}})} x_j > \sum_{j \in N_k(\bar{\mathbf{g}})} x_j = y_k$. This contradicts strict strategic complementarities.

Part (ii): Recall that we denote the Nash equilibrium effort level in the complete network, $\bar{\mathbf{g}}^c$, with $\mathbf{x}(\bar{\mathbf{g}}^c)$. Start by deleting a link $\bar{g}_{i,j}$ from $\bar{\mathbf{g}}^c$ and consider each player's best response to $\mathbf{x}(\bar{\mathbf{g}}^c)$ in $\bar{\mathbf{g}}^c - \bar{g}_{i,j}$. Agent i 's initial best response will be lower in $\bar{\mathbf{g}}^c - \bar{g}_{i,j}$ than in $\bar{\mathbf{g}}^c$, as $\sum_{j \in N_i(\bar{\mathbf{g}}^c - \bar{g}_{i,j})} x_j(\bar{\mathbf{g}}^c) < \sum_{j \in N_i(\bar{\mathbf{g}}^c)} x_j(\bar{\mathbf{g}}^c)$. Iterating on best responses, any agent l with $\bar{g}_{i,l} = 1$ decreases the effort level and in turn any agents sustaining links with l decrease their effort levels, and so forth. The effort level of each agent is a decreasing sequence of real numbers, which is bounded below by $\bar{x}(0)$. Effort levels therefore convergence to a new equilibrium in $\bar{\mathbf{g}}^c - \bar{g}_{i,j}$ with $x_l(\bar{\mathbf{g}}^c - \bar{g}_{i,j}) < x(\bar{\mathbf{g}}^c) \forall l \in N$. Note that any network $\bar{\mathbf{g}} \neq \bar{\mathbf{g}}^c$ can be obtained from $\bar{\mathbf{g}}^c$ by deleting a sequence of links. Effort levels are weakly decreasing at each step and strictly for the deletion of the first link in a complete network.

Part (iii): Part (iii) follows directly from the argument in Part (ii). *Q.E.D.*

Proof of Proposition 2. We start by showing that $\underline{\kappa} < \bar{\kappa}$. From part (ii) of Proposition 1 we know that $x(\bar{\mathbf{g}}^c) > x'_j(\bar{\mathbf{g}}^e + \bar{g}_{i,j})$ and, since v is (weakly) convex, $v_i(x'_j(\bar{\mathbf{g}}^e + \bar{g}_{i,j})) - v(0) < \frac{v((n-1)x(\bar{\mathbf{g}}^c)) - v(0)}{n-1}$ holds. We show the remaining part of the proof in two claims.

Claim 1: *If $\kappa < \underline{\kappa}$, then the unique PNE is such that the network is complete, while if $\kappa \geq \underline{\kappa}$, then there exists a PNE such that the network is empty.*

Denote with $x'_i(\bar{\mathbf{g}}^e + \bar{g}_{i,j}) = x'_j(\bar{\mathbf{g}}^e + \bar{g}_{i,j})$ the deviation effort levels when agents i and j create a link in the empty network $\bar{\mathbf{g}}^e$. Note that the deviation effort levels $x'_i(\bar{\mathbf{g}}^e + \bar{g}_{i,j}) = x'_j(\bar{\mathbf{g}}^e + \bar{g}_{i,j})$ are equal to the Nash equilibrium effort levels on a fixed network $\bar{\mathbf{g}}^e + \bar{g}_{i,j}$. If $\kappa < \underline{\kappa} = v_i(x'_j(\bar{\mathbf{g}}^e + \bar{g}_{i,j})) - v(0)$, then a pair of agents i and j finds it profitable to create the link $\bar{g}_{i,j}$ in $\bar{\mathbf{g}}^e$ and exert effort level $x'_i(\bar{\mathbf{g}}^e + \bar{g}_{i,j})$. Note that this is the least profitable link in any network. This follows from the (weak) convexity of the value function, strict strategic complementarities and the monotonicity property in effort levels when adding links to a network (Proposition 1). Therefore, for any $\kappa < \underline{\kappa}$, the unique PNE is the complete

network. If $\kappa \geq \underline{\kappa} = v_i(x'_j(\bar{\mathbf{g}}^e + \bar{g}_{i,j})) - v(0)$, then no pair of agents can profitably deviate in the empty network and a *PNE* exists such that the network is empty. *Q.E.D.*

Claim 2: *If $\kappa > \bar{\kappa}$, then the unique PNE is such that the network is empty, while if $\kappa \leq \bar{\kappa}$ then there exists a PNE such that the network is complete.*

Note first that the relevant deviation to consider in a complete network is an agent deleting all links. This follows directly from the convexity of the value function and that in any pairwise Nash equilibrium $x(\bar{\mathbf{g}}^c) = x_i(\bar{\mathbf{g}}^c) > x_i(\bar{\mathbf{g}}) \forall \bar{\mathbf{g}} \neq \bar{\mathbf{g}}^c$ (Proposition 1). For linking cost $\kappa \leq \bar{\kappa} = \frac{v((n-1)x(\bar{\mathbf{g}}^c)) - v(0)}{n-1}$ agents do not find it profitable to delete all their links and there therefore exists a *PNE* such that the network is complete. Conversely, for $\kappa > \bar{\kappa}$ agents in the complete network can profitably deviate by deleting all of their links. Next we show that if $\kappa > \bar{\kappa}$, then the unique *PNE* is the empty network. Assume to the contrary that $\kappa > \bar{\kappa}$ and there exists a *PNE* such that $\bar{\mathbf{g}} \notin \{\bar{\mathbf{g}}^c, \bar{\mathbf{g}}^e\}$. Note that then

$$\kappa > \frac{v((n-1)x(\bar{\mathbf{g}}^c)) - v(0)}{n-1} > \frac{v(\sum_{j \in N_i(\bar{\mathbf{g}})} x_j(\bar{\mathbf{g}})) - v(0)}{\eta_i(\bar{\mathbf{g}})}$$

holds $\forall i \in N$. Note that the second inequality follows from the (weak) convexity of the value function, that $x(\bar{\mathbf{g}}^c) > x_i(\bar{\mathbf{g}}) \forall \bar{\mathbf{g}} \neq \bar{\mathbf{g}}^c$ holds in any pairwise Nash equilibrium and $n-1 \geq \eta_i(\bar{\mathbf{g}})$. That is, if an agent finds it profitable to delete all links in $\bar{\mathbf{g}}^c$, then the agent also finds it profitable to delete all links in any network $\bar{\mathbf{g}} \notin \{\bar{\mathbf{g}}^c, \bar{\mathbf{g}}^e\}$. Therefore, the unique *PNE* is such that the network is empty. *Q.E.D.*

Proof of Theorem 1. We first provide four auxiliary lemmas, which we use to show that every *PNE* is a nested split graph. This will also yield Part (i) of Theorem 1. Note that when a pair of agents i and j deviates by creating a link, i.e. $\bar{g}_{i,j} = 0$ before the deviation and $\bar{g}_{i,j} = 1$ after the deviation, then, due to strategic complementarities, effort levels after the deviation are strictly larger than prior to it. That is, $x'_i > x_i(\bar{\mathbf{g}})$ and $x'_j > x_j(\bar{\mathbf{g}})$ hold.

Lemma 1: *In any PNE, $(\mathbf{x}, \bar{\mathbf{g}})$, if $\bar{g}_{i,l} = 1$, then $\bar{g}_{i,k} = 1$ for all agents k with $x_k \geq x_l$.*

Assume to the contrary of the above that $\bar{g}_{i,l} = 1$ and $\bar{g}_{i,k} = 0$ for some agent k with $x_k \geq x_l$. Note first that for $\bar{g}_{i,l} = 1$ to be part of a *PNE* it must be that $v(y_l) - v(y_l - x_l) \geq \kappa$. If the latter inequality holds, then, due to the (weak) convexity of the value function and $x'_k > x_l$, linking to any agent k with $x_k \geq x_l$ is also profitable for agent i . For $\bar{g}_{i,k} = 0$ to hold it must therefore be the case that agent k does not find it profitable to link to agent i . In the following we show that this cannot be the case. For $\bar{g}_{i,l} = 1$ to hold we must have $v(y_l) - v(y_l - x_i) \geq \kappa$. Note next that, due to $x_k \geq x_l$ and strict strategic complementarities, $y_k \geq y_l$ holds. We can now write $v(y_k + x'_i) - v(y_k) > v(y_k + x_i) - v(y_k) \geq v(y_l) - v(y_l - x_i) \geq \kappa$. The inequalities follow from the (weak) convexity of the value function, $x'_i > x_i$ and $y_k \geq y_l$. That is, if $\bar{g}_{i,l} = 1$, then agent i finds it profitable to link to any agent k with $x_k \geq x_l$, while any agent k finds it profitable to link to agent i . Therefore, in any *PNE*, $\bar{g}_{i,k} = 1$ for all agents k with $x_k \geq x_l$. *Q.E.D.*

Lemma 2: *In any PNE, $(\mathbf{x}, \bar{\mathbf{g}})$, $x_i = x_k \Leftrightarrow N_i(\bar{\mathbf{g}}) \setminus \{k\} = N_k(\bar{\mathbf{g}}) \setminus \{i\}$.*

First we show that $N_i(\bar{\mathbf{g}}) \setminus \{k\} = N_k(\bar{\mathbf{g}}) \setminus \{i\} \Rightarrow x_i = x_k$. If $\bar{g}_{i,k} = 0$, then $y_i = y_k$ and therefore $x_i = x_k$. Assume next that $\bar{g}_{i,k} = 1$ and, without loss of generality, that $x_i > x_k$. But then $y_i < y_k$ holds (since through the link $\bar{g}_{i,k} = 1$ agent i accesses x_k , while agent k accesses x_i) and we have reached a contradiction. Next we show that $x_i = x_k \Rightarrow N_i(\bar{\mathbf{g}}) \setminus \{k\} =$

$N_k(\bar{\mathbf{g}}) \setminus \{i\}$. Assume to the contrary that $x_i = x_k$ and $N_i(\bar{\mathbf{g}}) \setminus \{k\} \neq N_k(\bar{\mathbf{g}}) \setminus \{i\}$. There must then exist an agent l such that $k \in N_l(\bar{\mathbf{g}})$ and $i \notin N_l(\bar{\mathbf{g}})$. This contradicts Lemma 1. *Q.E.D.*

Lemma 3: *In any PNE, $(\mathbf{x}, \bar{\mathbf{g}})$, $x_i < x_k \Leftrightarrow N_i(\bar{\mathbf{g}}) \setminus \{k\} \subset N_k(\bar{\mathbf{g}}) \setminus \{i\}$.*

First we show that $N_i(\bar{\mathbf{g}}) \setminus \{k\} \subset N_k(\bar{\mathbf{g}}) \setminus \{i\} \Rightarrow x_i < x_k$. If $\bar{g}_{i,k} = 0$, then $y_k > y_i$ holds and therefore $x_i < x_k$ also holds, due to strict strategic complementarities. Assume next that $\bar{g}_{i,k} = 1$ and $x_i \geq x_k$. Then, for $N_i(\bar{\mathbf{g}}) \setminus \{k\} \subset N_k(\bar{\mathbf{g}}) \setminus \{i\}$ to hold, there must exist an agent m such that $m \in N_k(\bar{\mathbf{g}}) \setminus \{i\}$ and $m \notin N_i(\bar{\mathbf{g}}) \setminus \{k\}$. But then, since $x_i \geq x_k$, i and m must be connected by Lemma 1 and we have reached a contradiction. Next we show that $x_i < x_k \Rightarrow N_i(\bar{\mathbf{g}}) \setminus \{k\} \subset N_k(\bar{\mathbf{g}}) \setminus \{i\}$. Assume to the contrary that $x_i < x_k$ and $N_i(\bar{\mathbf{g}}) \setminus \{k\} \not\subset N_k(\bar{\mathbf{g}}) \setminus \{i\}$. We distinguish two subcases. Assume first that $x_i < x_k$ and $N_k(\bar{\mathbf{g}}) \setminus \{i\} = N_i(\bar{\mathbf{g}}) \setminus \{k\}$. This contradicts Lemma 2. Next, assume $x_i < x_k$ and $N_k(\bar{\mathbf{g}}) \setminus \{i\} \neq N_i(\bar{\mathbf{g}}) \setminus \{k\}$. There then exists an agent m such that $m \in N_i(\bar{\mathbf{g}}) \setminus \{k\}$ and $m \notin N_k(\bar{\mathbf{g}}) \setminus \{i\}$. But then, since $x_i < x_k$, i and m must be connected by Lemma 1 and we have reached a contradiction. *Q.E.D.*

Lemma 4: *In any PNE, $(\mathbf{x}, \bar{\mathbf{g}})$, $x_i \leq x_k \Leftrightarrow N_i(\bar{\mathbf{g}}) \setminus \{k\} \subseteq N_k(\bar{\mathbf{g}}) \setminus \{i\}$. Furthermore, $x_i < x_k \Leftrightarrow \eta_i(\bar{\mathbf{g}}) < \eta_k(\bar{\mathbf{g}})$, $x_i \leq x_k \Leftrightarrow \eta_i(\bar{\mathbf{g}}) \leq \eta_k(\bar{\mathbf{g}})$ and $x_i < x_k \Leftrightarrow \pi_i < \pi_k$.*

This first three equivalence relationships follow directly from the lemmas above. The fourth follows from strategic complementarities and an increasing value function.

In any PNE the network is a nested split graph.

In any PNE if $\bar{g}_{i,l} = 1$ and $\eta_k(\bar{\mathbf{g}}) \geq \eta_l(\bar{\mathbf{g}})$, then $x_k \geq x_l$ by Lemma 4 and $\bar{g}_{i,k} = 1$ by Lemma 1. That is, $\bar{\mathbf{g}}$ is a nested split graph. *Q.E.D.*

Part (ii): We show that if $\eta_i(\bar{\mathbf{g}}) \neq \eta_k(\bar{\mathbf{g}})$, then in any PNE $|\eta_i(\bar{\mathbf{g}}) - \eta_k(\bar{\mathbf{g}})| \neq 1$. Assume w.l.o.g. that $\eta_i(\bar{\mathbf{g}}) < \eta_k(\bar{\mathbf{g}})$ holds. From the above we then know that $x_i < x_k$ and $N_i(\bar{\mathbf{g}}) \setminus \{k\} \subset N_k(\bar{\mathbf{g}}) \setminus \{i\}$ also hold. Assume to the contrary that $\eta_i(\bar{\mathbf{g}}) + 1 = \eta_k(\bar{\mathbf{g}})$. There then exists a single agent l such that $\bar{g}_{l,k} = 1$ and $\bar{g}_{l,i} = 0$. Note that for $(\mathbf{x}, \bar{\mathbf{g}})$ to be a PNE network we must have that $v(y_k) - v(y_k - x_l) \geq \kappa$. We distinguish two cases. Assume first that $\bar{g}_{i,k} = 0$. Note that then $y_i + x_l = y_k$ and we can therefore write $v(y_i + x_l) - v(y_i) = v(y_k) - v(y_k - x_l)$. Consider a deviation such that $\bar{g}_{i,l} = 1$. By strategic complementarity $x'_l > x_l$ holds and the deviation is therefore profitable for agent i , since $v(y_i + x'_l) - v(y_i) > v(y_i + x_l) - v(y_i) = v(y_k) - v(y_k - x_l) \geq \kappa$. Likewise, from $y_i + x_l = y_k$, $x'_l > x_l$ and strategic complementarities we know that $x'_i > x_k$. Since the value function is (weakly) convex and given the link $\bar{g}_{l,k} = 1$, agent l also finds it profitable to link to agent i . Assume next that $\bar{g}_{i,k} = 1$. Note that then $y_i + x_l > y_k$ holds. We can now write $v(y_i + x'_l) - v(y_i) > v(y_i + x_l) - v(y_i) \geq v(y_k) - v(y_k - x_l) \geq \kappa$, where $v(y_i + x_l) - v(y_i) \geq v(y_k) - v(y_k - x_l)$ follows from the (weak) convexity of the value function and $y_i > y_k - x_l$. Therefore, agent i finds it profitable to link to agent l . From $y_i + x_l > y_k$, $x'_l > x_l$ and strategic complementarities we again know that $x'_i > x_k$. Since the value function is (weakly) convex and given the link $\bar{g}_{l,k} = 1$, agent l finds it profitable to link to agent i . *Q.E.D.*

Proof of Corollary 1. Assume $v''(y) = 0 \forall y$ and $(\mathbf{x}, \bar{\mathbf{g}})$ is a pairwise Nash equilibrium such that $\bar{\mathbf{g}}$ is a nested split graph that is neither complete, empty, or a dominant group network. Then there exists a pair of agents i and j such that $\eta_i(\bar{\mathbf{g}}) > \eta_j(\bar{\mathbf{g}}) > 0$. Since $\eta_i(\bar{\mathbf{g}}) > \eta_j(\bar{\mathbf{g}})$, there exist an agent k such that $\bar{g}_{i,k} = 1$ and $\bar{g}_{j,k} = 0$. Note further that, since $\eta_j(\bar{\mathbf{g}}) > 0$, there exists an agent l such that $\bar{g}_{j,l} = 1$ (were we allow for $l = i$). For $\bar{\mathbf{g}}$ to be a pairwise Nash equilibrium it must be the case that $v(y_i(\bar{\mathbf{g}})) - v(y_i(\bar{\mathbf{g}}) - x_k(\bar{\mathbf{g}})) \geq k$ and $v(y_l(\bar{\mathbf{g}})) - v(y_l(\bar{\mathbf{g}}) - x_j(\bar{\mathbf{g}})) \geq \kappa$. But then agents j and k find it profitable to create the link $\bar{g}_{j,k} = 1$, since $v(y_j(\bar{\mathbf{g}}) + x'_k) - v(y_j(\bar{\mathbf{g}})) > v(y_i(\bar{\mathbf{g}})) - v(y_i(\bar{\mathbf{g}}) - x_k(\bar{\mathbf{g}})) \geq \kappa$ and $v(y_k(\bar{\mathbf{g}}) + x'_j) - v(y_k(\bar{\mathbf{g}})) > v(y_l(\bar{\mathbf{g}})) - v(y_l(\bar{\mathbf{g}}) - x_j(\bar{\mathbf{g}})) \geq \kappa$. The inequalities follow directly from $v''(y) = 0 \forall y$ and $x'_k > x_k(\bar{\mathbf{g}})$ and $x'_j > x_j(\bar{\mathbf{g}})$. Note also that any group dominant network such that $\eta_i(\bar{\mathbf{g}}) = 1$ (i.e. a pair of connected agents, while the remaining agents are singletons) can not be a pairwise Nash equilibrium by Theorem 1. *Q.E.D.*

Proof of Corollary 2. Part (i): The statement follows directly from Proposition 2. For $\kappa < \underline{\kappa}$ the only *PNE* is such that the network is complete, while for $\kappa > \bar{\kappa}$ the only *PNE* is such that the network is empty. For any *PNE* such that $\bar{\mathbf{g}} \notin \{\bar{\mathbf{g}}^c, \bar{\mathbf{g}}^e\}$ this leaves only the region of $\kappa \in [\underline{\kappa}, \bar{\kappa}]$. We know from Proposition 2 that in this interval there always exists a *PNE* such that the network is complete and another *PNE* in which it is empty.

Part (ii): The statement follows directly from Proposition 1 and we omit the proof.

Part (iii): Assume that there exists a *PNE* with $\bar{\mathbf{g}} \notin \{\bar{\mathbf{g}}^c, \bar{\mathbf{g}}^e\}$. Since agents have the highest number of links in $\bar{\mathbf{g}}^c$, it is sufficient to show that the average payoff per link is highest in the *PNE* with $\bar{\mathbf{g}}^c$. From Proposition 1 we know that $x(\bar{\mathbf{g}}^c) = x_i(\bar{\mathbf{g}}^c) > x_i(\bar{\mathbf{g}}) \forall \bar{\mathbf{g}} \neq \bar{\mathbf{g}}^c$. For $\bar{\mathbf{g}} \notin \{\bar{\mathbf{g}}^c, \bar{\mathbf{g}}^e\}$ pick \tilde{x} such that $x(\bar{\mathbf{g}}^c) > \tilde{x} \geq x_i(\bar{\mathbf{g}}) \forall i \in N$. Since the value function is (weakly) convex we know that

$$\frac{v((n-1)x(\bar{\mathbf{g}}^c))}{n-1} > \frac{v((n-1)\tilde{x})}{n-1} \geq \frac{v(\sum_{j \in N_i(\bar{\mathbf{g}})} x_j)}{\eta_i(\bar{\mathbf{g}})}$$

holds $\forall i \in N$ and $\forall \bar{\mathbf{g}} \neq \bar{\mathbf{g}}^c$. Note further that the payoff of each agent in a *PNE* such that $\bar{\mathbf{g}} \notin \{\bar{\mathbf{g}}^c, \bar{\mathbf{g}}^e\}$ must be at least as large as in the *PNE* where the network is empty. Otherwise, a profitable deviation exists. Therefore, individual payoffs are maximal in the *PNE* such that the network is complete. *Q.E.D.*

Proof of Proposition 3. Consider first any link such that $\bar{g}_{i,j} = 0$ and $\bar{g}_{i,j}'' = 0$. From the monotonicity properties of Proposition 1 in the main part of the paper and $\bar{\mathbf{g}}' \subset \bar{\mathbf{g}} \subset \bar{\mathbf{g}}''$ it follows immediately that if $\bar{g}_{i,j}'' = 0$, then i and j can also not profitably deviate by creating the link $\bar{g}_{i,j} = 1$ in $\bar{\mathbf{g}}$ and therefore $\bar{g}_{i,j} = 0$. Next consider any link such that $\bar{g}_{i,j} = 1$ and $\bar{g}_{i,j}' = 1$. Again from Proposition 1 and $\bar{\mathbf{g}}' \subset \bar{\mathbf{g}} \subset \bar{\mathbf{g}}''$ it follows directly that there is no profitable deviation in which an agent deletes any links in $\bar{\mathbf{g}}$ that are already in place in $\bar{\mathbf{g}}'$. Third, consider any link such that $\bar{g}_{i,j} = 1$ and $\bar{g}_{i,j}' = 0$. We assume that there exists a link $\bar{g}_{k,l}' = 1$ in $\bar{\mathbf{g}}'$ such that $\bar{g}_{i,j} \succeq \bar{g}_{k,l}'$. That is, there exists a link in $\bar{\mathbf{g}}$ such that (without loss of generality) $\eta_i(\bar{\mathbf{g}}) \geq \eta_k(\bar{\mathbf{g}}')$ and $\eta_j(\bar{\mathbf{g}}) \geq \eta_l(\bar{\mathbf{g}}')$. From Proposition 1 in the main part of the paper, $\bar{\mathbf{g}}' \subset \bar{\mathbf{g}}$ and since there is at most one connected component in any pairwise Nash equilibrium, we know that $x_i(\bar{\mathbf{g}}) \geq x_i(\bar{\mathbf{g}}') \forall i$ and $x_i(\bar{\mathbf{g}}) > x_i(\bar{\mathbf{g}}') \forall i : \eta_i(\bar{\mathbf{g}}) \geq 1$. Note that in a nested split graph an agent i with η_i links is linked to the η_i agents with the highest number of links. From $x_i(\bar{\mathbf{g}}) < x_j(\bar{\mathbf{g}}) \Leftrightarrow \eta_i(\bar{\mathbf{g}}) < \eta_j(\bar{\mathbf{g}})$ (Theorem 1) we also know that these are also the η_i agents with the highest effort level. Therefore, $y_i(\bar{\mathbf{g}}) > y_k(\bar{\mathbf{g}}')$ and $y_j(\bar{\mathbf{g}}) > y_l(\bar{\mathbf{g}}')$ holds and, from strategic complementarities, $x_i(\bar{\mathbf{g}}) > x_k(\bar{\mathbf{g}}')$ and $x_j(\bar{\mathbf{g}}) > x_l(\bar{\mathbf{g}}')$ also holds. From the (weak) convexity of the value function it then follows that if neither k nor l find it

profitable to delete their link in $\bar{\mathbf{g}}'$, then neither i nor j find it profitable to delete their link in $\bar{\mathbf{g}}$. The argument for any link such that $\bar{g}_{i,j} = 0$ and $\bar{g}'_{i,j} = 1$ is analogous. If there exists a link $\bar{g}'_{k,l} = 0$ such that $\bar{g}'_{k,l} \succsim \bar{g}'_{i,j}$, then, given that k and l do not find it profitable to create the link in $\bar{\mathbf{g}}''$, neither i nor j find it profitable to create a link in $\bar{\mathbf{g}}''$. *Q.E.D.*

Proof of Proposition 4. We distinguish two cases. First we consider the case of payoff functions, such that all best response functions are strictly concave. Note that Kennan's (2001) result is formulated in terms of strictly concave functions from \mathbb{R}^n to \mathbb{R}^n and it is straightforward to adapt the proof to account for heterogeneous best response functions. For linear best response functions we apply Banach's fixed point theorem.³¹ Define the function $f_{\bar{\mathbf{g}},\boldsymbol{\theta}} : X^n \rightarrow X^n$ as

$$f_{\bar{\mathbf{g}},\boldsymbol{\theta}}(\mathbf{x}) = \begin{pmatrix} \bar{x}_1 \left(\sum_{j \in N_1(\bar{\mathbf{g}})} x_j, \boldsymbol{\theta}_1 \right) \\ \vdots \\ \bar{x}_n \left(\sum_{j \in N_n(\bar{\mathbf{g}})} x_j, \boldsymbol{\theta}_n \right) \end{pmatrix},$$

where we allow for individual best response functions of the form $\bar{x}_i \left(\sum_{j \in N_i(\bar{\mathbf{g}})} x_j, \boldsymbol{\theta}_i \right) = \frac{1}{\beta_i} + \frac{\lambda_i}{\beta_i} \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$, so that we can write $\boldsymbol{\theta}_i = (\lambda_i, \beta_i)$, and we assume $\frac{\lambda_i}{\beta_i} < \frac{1}{n-1} \forall i \in N$. Note that writing the best response function in this form is without loss of generality. For any pair of vectors $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ define $d(\mathbf{x}, \mathbf{x}')$ as $d(\mathbf{x}, \mathbf{x}') = \sum_{i \in N} |x_i - x'_i|$. To show that a unique Nash equilibrium exists, it is then sufficient to show that $f_{\bar{\mathbf{g}},\boldsymbol{\theta}}$ is a contraction mapping, i.e. there exists a $q \in (0, 1)$ such that $\forall \mathbf{x}, \mathbf{x}' \in X^n$, $d(f_{\bar{\mathbf{g}},\boldsymbol{\theta}}(\mathbf{x}), f_{\bar{\mathbf{g}},\boldsymbol{\theta}}(\mathbf{x}')) \leq q \cdot d(\mathbf{x}, \mathbf{x}')$. Define $\tilde{q} = \max\{\frac{\lambda_1}{\beta_1}, \dots, \frac{\lambda_n}{\beta_n}\}$ and set $q = (n-1)\tilde{q}$. Note that $q \in (0, 1)$ since $0 < \frac{\lambda_i}{\beta_i} < \frac{1}{n-1} \forall i \in N$. We first show that $f_{\bar{\mathbf{g}},\boldsymbol{\theta}}$ is a contraction mapping if $\bar{\mathbf{g}} = \bar{\mathbf{g}}^c$. If $\bar{\mathbf{g}} = \bar{\mathbf{g}}^c$, then $d(f_{\bar{\mathbf{g}},\boldsymbol{\theta}}(\mathbf{x}), f_{\bar{\mathbf{g}},\boldsymbol{\theta}}(\mathbf{x}'))$ can be written as

$$\begin{aligned} d(f_{\bar{\mathbf{g}}^c,\boldsymbol{\theta}}(\mathbf{x}), f_{\bar{\mathbf{g}}^c,\boldsymbol{\theta}}(\mathbf{x}')) &= \sum_{i \in N} \left| \frac{\lambda_i}{\beta_i} \sum_{j \neq i} (x_j - x'_j) \right| \leq \sum_{i \in N} \tilde{q} \left| \sum_{j \neq i} (x_j - x'_j) \right| \\ &= (n-1)\tilde{q} \sum_{i \in N} |x_i - x'_i| = q \sum_{i \in N} |x_i - x'_i| \\ &< \sum_{i \in N} |x_i - x'_i| = d(\mathbf{x}, \mathbf{x}'). \end{aligned}$$

Note that if $\bar{\mathbf{g}} \neq \bar{\mathbf{g}}^c$, then some of the difference terms drop out relative to $d(f_{\bar{\mathbf{g}}^c,\boldsymbol{\theta}}(\mathbf{x}), f_{\bar{\mathbf{g}}^c,\boldsymbol{\theta}}(\mathbf{x}'))$ and $d(f_{\bar{\mathbf{g}},\boldsymbol{\theta}}(\mathbf{x}), f_{\bar{\mathbf{g}},\boldsymbol{\theta}}(\mathbf{x}')) \leq d(f_{\bar{\mathbf{g}}^c,\boldsymbol{\theta}}(\mathbf{x}), f_{\bar{\mathbf{g}}^c,\boldsymbol{\theta}}(\mathbf{x}'))$, while $d(\mathbf{x}, \mathbf{x}')$ is independent of $\bar{\mathbf{g}}$. Therefore, $f_{\bar{\mathbf{g}},\boldsymbol{\theta}}$ is a contraction mapping and there exists a unique Nash equilibrium. The second and third part of the statement directly follow from the argument in Proposition 1. *Q.E.D.*

For $X \subseteq \mathbb{R}^l$, define $K(X) = \{K \subseteq X : K \neq \emptyset \text{ is compact}\}$. For a set X , denote $X^n = X_1 \times \dots \times X_n$.

Definition 4: (Theorem of the Maximum). For $\Theta \subseteq \mathbb{R}^k$, $X \subseteq \mathbb{R}$, $\Psi : \Theta \rightarrow K(X)$ a correspondence and $\pi : X \times \Theta \rightarrow \mathbb{R}$, define the value function $\boldsymbol{\theta} \rightarrow v(\boldsymbol{\theta}) = \max_{x \in \Psi(\boldsymbol{\theta})} \pi(x, \boldsymbol{\theta})$, and the argmax correspondence $\boldsymbol{\theta} \rightarrow \bar{x}(\boldsymbol{\theta}) = \{x \in \Psi(\boldsymbol{\theta}) : \forall x' \in \Psi(\boldsymbol{\theta}), \pi(x, \boldsymbol{\theta}) \geq \pi(x', \boldsymbol{\theta})\}$.

³¹If $f : M \rightarrow M$ is a contraction mapping and (M, d) is a complete metric space, then f has a unique fixed point, \mathbf{x} . Note further that \mathbb{R}^n is a complete metric space with the absolute value metric. See Corbae et al. (2009), p. 121.

If π is (jointly) continuous and $\Psi(\boldsymbol{\theta})$ is continuous then (i) $v(\cdot)$ is continuous and (ii) $\bar{x}(\cdot)$ is hemicontinuous. Furthermore, if $\bar{x}(\cdot)$ is always singleton valued, then $\boldsymbol{\theta} \rightarrow \bar{\mathbf{x}}(\boldsymbol{\theta})$ is continuous.³²

Definition 5: (Implicit Function Theorem). Let $f_{\bar{\mathbf{g}}} : \mathbb{R}^{n+nk} \supseteq X^n \times \Theta^n \rightarrow \mathbb{R}^n$ be a continuously differentiable function on an open set $X^n \times \Theta^n$. Consider the system of equations $f_{\bar{\mathbf{g}}}(\mathbf{x}, \boldsymbol{\theta}) = \mathbf{0}$ and assume it has a solution at $\mathbf{x}^0 \in X^n$ for given parameter values $\boldsymbol{\theta}^0 \in \Theta^n$. If the determinant of the Jacobian of endogenous variables is not zero at $(\mathbf{x}^0, \boldsymbol{\theta}^0)$, that is, if $|J(\mathbf{x}^0, \boldsymbol{\theta}^0)| = |D_{\mathbf{x}} f_{\bar{\mathbf{g}}}(\mathbf{x}^0, \boldsymbol{\theta}^0)| \neq 0$ then (i) there exist open sets U in \mathbb{R}^{n+nk} and $U_{\boldsymbol{\theta}}$ in \mathbb{R}^{nk} with $(\mathbf{x}^0, \boldsymbol{\theta}^0) \subseteq U$ and $\boldsymbol{\theta}^0 \subseteq U_{\boldsymbol{\theta}}$ such that for each $\boldsymbol{\theta}$ in $U_{\boldsymbol{\theta}}$ there exists a unique $\mathbf{x}_{\boldsymbol{\theta}}$ such that $(\mathbf{x}_{\boldsymbol{\theta}}, \boldsymbol{\theta}) \in U$ and $f_{\bar{\mathbf{g}}}(\mathbf{x}_{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \mathbf{0}$. That is, the correspondence from $U_{\boldsymbol{\theta}}$ to X^n defined by $\mathbf{x}(\boldsymbol{\theta}) = \mathbf{x}_{\boldsymbol{\theta}}$ is a well-defined function when restricted to U . (ii) The solution function $\mathbf{x}(\cdot) : U_{\boldsymbol{\theta}} \rightarrow \mathbb{R}^n$ is continuously differentiable. (iii) If $f_{\bar{\mathbf{g}}}$ is C^k , so is $\mathbf{x}(\cdot)$.³³

Proof of Proposition 5. We start by showing that best response and value functions change continuously in the parameters of the payoff function. Assume $(\mathbf{x}, \bar{\mathbf{g}})$ is a *SPNE* for the homogeneous agent case, i.e. $\boldsymbol{\theta}_j = \boldsymbol{\theta}_i \forall i, j \in N$. Next, write an agent i 's payoff function as $\pi_i(x_i, y_i, \boldsymbol{\theta}_i)$. Recall that $\boldsymbol{\theta}_i \in \Theta$, Θ is open and $\Theta \subseteq \mathbb{R}^k$. Define $\tilde{\boldsymbol{\theta}}_i = (y_i, \boldsymbol{\theta}_i)$ with $\tilde{\boldsymbol{\theta}}_i \in \tilde{\Theta}$ and $\tilde{\Theta} \subseteq \mathbb{R}^{k+1}$. We can now write $\pi_i(x_i, \tilde{\boldsymbol{\theta}}_i)$. Similarly, we can write for agent i 's value and best response function $v_i(\tilde{\boldsymbol{\theta}}_i)$ and $\bar{x}_i(\tilde{\boldsymbol{\theta}}_i)$, respectively. Define $\tilde{X} = [0, M]$ and assume M sufficiently large, so that the vector of effort levels of the *SPNE* with homogeneous agents, \mathbf{x} , is in the interior of $[0, M]^n$. Furthermore, assume $\Psi(\tilde{\boldsymbol{\theta}}_i) = [0, M] \forall \tilde{\boldsymbol{\theta}}_i \in \tilde{\Theta}$. Note that $\Psi(\tilde{\boldsymbol{\theta}}_i) = [0, M] \forall \tilde{\boldsymbol{\theta}}_i \in \tilde{\Theta}$ imposes a restriction on agents' strategies, since $x_i \in [0, M]$ must hold $\forall i \in N$. However, $(\mathbf{x}, \bar{\mathbf{g}})$ is also a *SPNE* in the presence of the restriction, since *SPNE* effort levels \mathbf{x} are feasible when $\Psi(\tilde{\boldsymbol{\theta}}_i) = [0, M] \forall \tilde{\boldsymbol{\theta}}_i \in \tilde{\Theta}$, while the set of available deviations is restricted. Note further that $\Psi(\tilde{\boldsymbol{\theta}}_i)$ is continuous. Moreover, π is (jointly) continuous and $\bar{x}_i(\tilde{\boldsymbol{\theta}}_i)$ is singleton valued by assumption. From the Theorem of the Maximum (Definition 4) it then follows directly that $v_i(\tilde{\boldsymbol{\theta}}_i)$ and $\bar{x}_i(\tilde{\boldsymbol{\theta}}_i)$ are continuous in $\tilde{\boldsymbol{\theta}}_i$. Next, define $\tilde{X}' = (0, M)$ and the function $f_{\bar{\mathbf{g}}} : \tilde{X}'^n \times \Theta^n \rightarrow \mathbb{R}^n$ as

$$f_{\bar{\mathbf{g}}}(\mathbf{x}, \boldsymbol{\theta}) = \begin{pmatrix} \partial \pi_1(x_1, \sum_{j \in N_1(\bar{\mathbf{g}})} x_j, \boldsymbol{\theta}_1) / \partial x_1 \\ \vdots \\ \partial \pi_n(x_n, \sum_{j \in N_1(\bar{\mathbf{g}})} x_j, \boldsymbol{\theta}_n) / \partial x_n \end{pmatrix}.$$

Given our assumptions on π , $f_{\bar{\mathbf{g}}}(\mathbf{x}, \boldsymbol{\theta})$ is continuously differentiable and, since \mathbf{x} is assumed to be the vector of *SPNE* effort levels with homogeneous agents, $f_{\bar{\mathbf{g}}}(\mathbf{x}, \boldsymbol{\theta}) = \mathbf{0}$ holds for the particular $\boldsymbol{\theta}$ considered for this case (i.e. where $\boldsymbol{\theta}$ is such that $\boldsymbol{\theta}_i = \boldsymbol{\theta}_j \forall i, j \in N$). We can now apply the Implicit Function Theorem. That is, if the Jacobian of $f_{\bar{\mathbf{g}}}(\mathbf{x}, \boldsymbol{\theta})$ is invertible, then the vector of equilibrium effort levels on a fixed network $\bar{\mathbf{g}}$, which we denote by $\mathbf{x}_{\bar{\mathbf{g}}}(\boldsymbol{\theta})$, is continuous in $\boldsymbol{\theta}$. Note that assuming $\tilde{X}' = (0, M)$ is analytically innocuous, since the vector of *SPNE* effort levels for the homogeneous agent case, \mathbf{x} , is a solution in the interior of $[0, M]^n$ and therefore in the interior of $(0, M)^n$. From Sard's Theorem we know that the set of critical points of a sufficiently smooth function has Lebesgue measure zero.³⁴ We

³²See Corbae et al. (2009), p. 151.

³³See de la Fuente (2009), p. 210.

³⁴Sard's Theorem reads as follows (see de la Fuente (2009), p. 214.). Let $g : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R}^m$ (X open) be a C^r function with $r > \max\{0, n - m\}$ and let C_f be the set of critical points of g . Then $g(C_f)$ has Lebesgue measure zero.

assume π to be smooth and therefore the property that $(\mathbf{x}, \boldsymbol{\theta})$ is a regular point of $f_{\bar{\mathbf{g}}}$ (and that the Jacobian of $f_{\bar{\mathbf{g}}}$ is invertible) is generic. That is, generically $\mathbf{x}_{\bar{\mathbf{g}}}(\boldsymbol{\theta})$ is continuous in $\boldsymbol{\theta}$. Since agents' best response functions, $\bar{x}_i(\tilde{\boldsymbol{\theta}}_i)$, and value functions, $v_i(\tilde{\boldsymbol{\theta}}_i)$, are continuous in $\tilde{\boldsymbol{\theta}}_i$, gross equilibrium payoffs are also continuous in $\boldsymbol{\theta}$. Note next that linking cost are assumed to be linear and agents' equilibrium payoffs are therefore also continuous in $\boldsymbol{\kappa}$. That is, equilibrium payoffs are generically continuous in $\boldsymbol{\theta}$ and $\boldsymbol{\kappa}$. Next we consider two types of deviations. First, deviations such that the network is altered. Again, since agents' best response functions, $\bar{x}_i(\tilde{\boldsymbol{\theta}}_i)$, and value functions, $v_i(\tilde{\boldsymbol{\theta}}_i)$, are continuous in $\tilde{\boldsymbol{\theta}}_i$, deviation effort levels and deviation payoffs are continuous in $\boldsymbol{\theta}$ and $\boldsymbol{\kappa}$. Second, deviations such that the network is not altered. Recall that *SPNE* allows for unilateral deviations such that the undirected network $\bar{\mathbf{g}}$ is not altered and payoffs after proposed deviation remain the same. Given *SPNE* link announcements \mathbf{g} , all such deviations also do not alter the undirected network $\bar{\mathbf{g}}$ when introducing heterogeneity in $\boldsymbol{\theta}$ and $\boldsymbol{\kappa}$ and therefore no such deviation is profitable. That is, if $(\mathbf{x}, \bar{\mathbf{g}})$ is a *SPNE* for the homogeneous agent case and sufficiently small heterogeneity is introduced in $\boldsymbol{\theta}$ and $\boldsymbol{\kappa}$, then there generically exists a *SPNE*, $(\tilde{\mathbf{x}}, \bar{\mathbf{g}})$, where $\tilde{\mathbf{x}} = \mathbf{x}_{\bar{\mathbf{g}}}(\boldsymbol{\theta})$. *Q.E.D.*

Definition 6: Let $N_{j \setminus i}(\bar{\mathbf{g}}) = \{k \in N \setminus \{i, j\} : \bar{g}_{j,k} = 1 \wedge \bar{g}_{i,k} = 0\}$ be the set of agents k that are connected to j but not to i in network $\bar{\mathbf{g}}$. An $N_{(j,i)}$ -switch then reallocates all links between j and all agents $k \in N_{j \setminus i}(\bar{\mathbf{g}})$ to links between i and all agents $k \in N_{j \setminus i}(\bar{\mathbf{g}})$. The network after an $N_{(j,i)}$ -switch, which we denote in the following as $\bar{\mathbf{g}}'$, is given by $\bar{\mathbf{g}}' = \bar{\mathbf{g}} - \sum_{k \in N_{j \setminus i}(\bar{\mathbf{g}})} \bar{g}_{j,k} + \sum_{k \in N_{j \setminus i}(\bar{\mathbf{g}})} \bar{g}_{i,k}$.

Proof of Theorem 2. Part (i) We will use Part (i) to prove that efficient networks are nested split graphs. We start by showing that $\hat{y}_i(\hat{\mathbf{g}}) < \hat{y}_j(\hat{\mathbf{g}}) \Rightarrow \hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}})$. Assume first and without loss of generality that $\exists i, j \in N : \hat{y}_i(\hat{\mathbf{g}}) < \hat{y}_j(\hat{\mathbf{g}})$ and $\hat{x}_i(\hat{\mathbf{g}}) > \hat{x}_j(\hat{\mathbf{g}})$. There then exists an agent k such that $\hat{g}_{k,j} = 1$ and $\hat{g}_{k,i} = 0$. Consider now a $N_{(j,i)}$ -switch (Definition 6) and denote the resulting network with $\bar{\mathbf{g}}'$. All agents k , whose links are switched from j to i in $\bar{\mathbf{g}}'$, obtain strictly higher gross payoffs since $\hat{x}_i(\hat{\mathbf{g}}) > \hat{x}_j(\hat{\mathbf{g}})$. Agent i 's gross payoffs from linking to agents k in $\bar{\mathbf{g}}'$ is strictly larger than for j in $\hat{\mathbf{g}}$ since $\frac{\partial^2 \Pi(x,y)}{\partial x \partial y} > 0$ and $y_i(\bar{\mathbf{g}}') \geq \hat{y}_j(\hat{\mathbf{g}})$. Gross payoffs for all other agents remain the same. Note that $\eta(\hat{\mathbf{g}}) = \eta(\bar{\mathbf{g}}')$ and $\Phi(\hat{\mathbf{g}}) = \Phi(\bar{\mathbf{g}}')$ hold and therefore $W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \bar{\mathbf{g}}') > W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}})$. Assume next that $\exists i, j : \hat{y}_i(\hat{\mathbf{g}}) < \hat{y}_j(\hat{\mathbf{g}})$ and $\hat{x}_i(\hat{\mathbf{g}}) = \hat{x}_j(\hat{\mathbf{g}})$. We distinguish two cases. If there exists an agent k such that $\hat{g}_{k,i} = 1$ and $\hat{g}_{k,j} = 0$, then rewiring the link such that $\bar{g}'_{k,i} = 0$ and $\bar{g}'_{k,j} = 1$ increases the total sum of payoffs since $\hat{y}_i(\hat{\mathbf{g}}) < \hat{y}_j(\hat{\mathbf{g}})$ and $\frac{\partial^2 \Pi(x,y)}{\partial x \partial y} > 0$. If $N_i(\hat{\mathbf{g}}) \setminus \{j\} \subset N_j(\hat{\mathbf{g}}) \setminus \{i\}$, then from $\hat{x}_i(\hat{\mathbf{g}}) = \hat{x}_j(\hat{\mathbf{g}})$ and $\frac{\partial^2 \Pi(x,y)}{\partial x \partial y} > 0$ we know that at least one first order condition does not hold and proposed configuration is not efficient.

Next we show that $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}}) \Rightarrow \hat{y}_i(\hat{\mathbf{g}}) < \hat{y}_j(\hat{\mathbf{g}})$. Assume first that $\exists i, j \in N : \hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}})$ and $\hat{y}_i(\hat{\mathbf{g}}) > \hat{y}_j(\hat{\mathbf{g}})$. We showed above that this cannot be part of an efficient configuration. Assume next that $\exists i, j \in N : \hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}})$ and $\hat{y}_i(\hat{\mathbf{g}}) = \hat{y}_j(\hat{\mathbf{g}})$. Then in an efficient configuration either $N_i(\hat{\mathbf{g}}) \setminus \{j\} = N_j(\hat{\mathbf{g}}) \setminus \{i\}$ or $N_i(\hat{\mathbf{g}}) \setminus \{j\} \subset N_j(\hat{\mathbf{g}}) \setminus \{i\}$ must hold. To see this, note that otherwise there exists an agent k such that $\hat{g}_{i,k} = 1$ and $\hat{g}_{j,k} = 0$. Since we assume $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}})$ and $\hat{y}_i(\hat{\mathbf{g}}) = \hat{y}_j(\hat{\mathbf{g}})$, rewiring the link such that $\bar{g}'_{i,k} = 0$ and $\bar{g}'_{j,k} = 1$ again increases the total sum of gross payoffs (while keeping network cost constant). We first consider the case when $N_i(\hat{\mathbf{g}}) \setminus \{j\} = N_j(\hat{\mathbf{g}}) \setminus \{i\}$. Then i and j cannot be connected as otherwise $\hat{y}_i(\hat{\mathbf{g}}) > \hat{y}_j(\hat{\mathbf{g}})$. Assume therefore that $\hat{g}_{i,j} = 0$. From $\frac{\partial^2 \Pi(x,y)}{\partial^2 y} \geq 0$ and $\frac{\partial^2 \Pi(x,y)}{\partial x \partial y} > 0$ we know that the first order conditions for $W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}})$ cannot hold for

both $\hat{x}_i(\hat{\mathbf{g}})$ and $\hat{x}_j(\hat{\mathbf{g}})$ when $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}})$ and $N_i(\hat{\mathbf{g}}) \setminus \{j\} = N_j(\hat{\mathbf{g}}) \setminus \{i\}$ hold and proposed configuration can not be efficient. Consider next the case when $N_i(\hat{\mathbf{g}}) \setminus \{j\} \subset N_j(\hat{\mathbf{g}}) \setminus \{i\}$. Note that then i and j must be connected as otherwise $\hat{y}_j(\hat{\mathbf{g}}) > \hat{y}_i(\hat{\mathbf{g}})$. Furthermore, for $\hat{y}_i(\hat{\mathbf{g}}) = \hat{y}_j(\hat{\mathbf{g}})$ to hold, it must be the case that the sum of effort levels of all agents $k \in N_{j \setminus i}(\hat{\mathbf{g}})$ is equal to $\hat{x}_j(\hat{\mathbf{g}}) - \hat{x}_i(\hat{\mathbf{g}})$. This implies that $\hat{x}_j(\hat{\mathbf{g}}) > \hat{x}_k(\hat{\mathbf{g}})$ holds $\forall k \in N_{j \setminus i}(\hat{\mathbf{g}})$. Denote by $W(\hat{x}_i(\hat{\mathbf{g}}), \hat{\mathbf{x}}_{-j}(\hat{\mathbf{g}}), \hat{\mathbf{g}})$ an alternative configuration (with the same $\hat{\mathbf{g}}$), where agent j 's effort level is set to the effort level $\hat{x}_i(\hat{\mathbf{g}})$. Denote by $W(\hat{x}_j(\hat{\mathbf{g}}), \hat{\mathbf{x}}_{-i}(\hat{\mathbf{g}}), \hat{\mathbf{g}})$ a further alternative configuration (again with the same $\hat{\mathbf{g}}$), where i 's effort level is set to the effort level $\hat{x}_j(\hat{\mathbf{g}})$. From $\frac{\partial^2 \Pi(x, y)}{\partial^2 y} \geq 0$ and $\frac{\partial^2 \Pi(x, y)}{\partial x \partial y} > 0$, together with $\hat{x}_j(\hat{\mathbf{g}}) > \hat{x}_k(\hat{\mathbf{g}}) \forall k \in N_{j \setminus i}(\hat{\mathbf{g}})$, it then follows immediately that $W(\hat{x}_j(\hat{\mathbf{g}}), \hat{\mathbf{x}}_{-i}(\hat{\mathbf{g}}), \hat{\mathbf{g}}) - W(\hat{\mathbf{g}}, \hat{\mathbf{x}}(\hat{\mathbf{g}})) > W(\hat{\mathbf{g}}, \hat{\mathbf{x}}(\hat{\mathbf{g}})) - W(\hat{x}_i(\hat{\mathbf{g}}), \hat{\mathbf{x}}_{-j}(\hat{\mathbf{g}}), \hat{\mathbf{g}})$ and $W(\hat{\mathbf{g}}, \hat{\mathbf{x}}(\hat{\mathbf{g}}))$ is therefore not efficient. That is, $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}}) \Leftrightarrow \hat{y}_i(\hat{\mathbf{g}}) < \hat{y}_j(\hat{\mathbf{g}})$ holds in any efficient configuration.

Next we show that $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}}) \Leftrightarrow \eta_i(\hat{\mathbf{g}}) < \eta_j(\hat{\mathbf{g}})$ holds. We start by showing that $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}}) \Rightarrow N_i(\hat{\mathbf{g}}) \setminus \{j\} \subset N_j(\hat{\mathbf{g}}) \setminus \{i\}$. Assume to the contrary that $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}})$ and $N_i(\hat{\mathbf{g}}) \setminus \{j\} \not\subset N_j(\hat{\mathbf{g}}) \setminus \{i\}$ holds. We distinguish two subcases. We consider first $N_i(\hat{\mathbf{g}}) \setminus \{j\} \neq N_j(\hat{\mathbf{g}}) \setminus \{i\}$. There then exists an agent k such that $\hat{g}_{k,i} = 1$ and $\hat{g}_{k,j} = 0$. Recall that we have shown that $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}}) \Leftrightarrow \hat{y}_i(\hat{\mathbf{g}}) < \hat{y}_j(\hat{\mathbf{g}})$ holds. As shown above, total payoffs increase when rewiring all links between i and $k \in N_{i \setminus j}(\hat{\mathbf{g}})$ to j and $k \in N_{i \setminus j}(\hat{\mathbf{g}})$. Assume next that $N_i(\hat{\mathbf{g}}) \setminus \{j\} = N_j(\hat{\mathbf{g}}) \setminus \{i\}$. If $\hat{g}_{i,j} = 1$, then $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}})$ implies $\hat{y}_i(\hat{\mathbf{g}}) > \hat{y}_j(\hat{\mathbf{g}})$ and we have reached a contradiction. If $\hat{g}_{i,j} = 0$, then $\hat{y}_i(\hat{\mathbf{g}}) = \hat{y}_j(\hat{\mathbf{g}})$ and again we have reached a contradiction. Next we show that $N_i(\hat{\mathbf{g}}) \setminus \{j\} \subset N_j(\hat{\mathbf{g}}) \setminus \{i\} \Rightarrow \hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}})$. Assume to the contrary that $N_i(\hat{\mathbf{g}}) \setminus \{j\} \subset N_j(\hat{\mathbf{g}}) \setminus \{i\}$ and $\hat{x}_i(\hat{\mathbf{g}}) \geq \hat{x}_j(\hat{\mathbf{g}})$ holds. Note that $\hat{x}_i(\hat{\mathbf{g}}) \geq \hat{x}_j(\hat{\mathbf{g}})$ and $N_i(\hat{\mathbf{g}}) \setminus \{j\} \subset N_j(\hat{\mathbf{g}}) \setminus \{i\}$ imply $\hat{y}_j(\hat{\mathbf{g}}) > \hat{y}_i(\hat{\mathbf{g}})$ and we have reached a contradiction. Therefore, $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}}) \Leftrightarrow N_i(\hat{\mathbf{g}}) \setminus \{j\} \subset N_j(\hat{\mathbf{g}}) \setminus \{i\}$ holds. Finally, $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}}) \Leftrightarrow N_i(\hat{\mathbf{g}}) \setminus \{j\} \subset N_j(\hat{\mathbf{g}}) \setminus \{i\}$ implies $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}}) \Leftrightarrow \eta_i(\hat{\mathbf{g}}) > \eta_j(\hat{\mathbf{g}})$.

We are now in the position to show that all efficient networks are nested split graphs. Recall our definition: If $\bar{g}_{i,j} = 1$ and $\eta_k(\bar{\mathbf{g}}) \geq \eta_l(\bar{\mathbf{g}})$, then $\bar{g}_{i,k} = 1$. Assume to the contrary that $\hat{g}_{i,l} = 1$ and $\eta_k(\hat{\mathbf{g}}) \geq \eta_l(\hat{\mathbf{g}})$, but $\hat{g}_{i,k} = 0$. If $\eta_k(\hat{\mathbf{g}}) > \eta_l(\hat{\mathbf{g}})$, then $\hat{x}_k(\hat{\mathbf{g}}) > \hat{x}_l(\hat{\mathbf{g}})$ and $\hat{y}_k(\hat{\mathbf{g}}) > \hat{y}_l(\hat{\mathbf{g}})$ also hold. From the above we know that rewiring the link $\hat{g}_{i,l} = 1$ such that $\hat{g}_{i,l} = 0$ and $\hat{g}_{i,k} = 1$ strictly increases the sum of payoffs. Assume next that $\eta_k(\hat{\mathbf{g}}) = \eta_l(\hat{\mathbf{g}})$. From the above we know that then $\hat{x}_k(\hat{\mathbf{g}}) = \hat{x}_l(\hat{\mathbf{g}})$ holds and therefore $N_k(\hat{\mathbf{g}}) \setminus \{l\} = N_l(\hat{\mathbf{g}}) \setminus \{k\}$, i.e. $\hat{g}_{i,k} = 1$. That is, any efficient configuration must be such that the network is a nested split graph.

Part (ii) We consider the case when $\Phi''(\eta) = 0 \forall \eta$. The case when $\Phi''(\eta) < 0 \forall \eta$ follows immediately. Assume there exists an efficient configuration such that the network is a nested split graph that is neither empty nor complete. Then there exists a pair of agents i and j such that $\eta_i(\hat{\mathbf{g}}) < \eta_j(\hat{\mathbf{g}})$. From the above we know that then also $\hat{x}_i(\hat{\mathbf{g}}) < \hat{x}_j(\hat{\mathbf{g}})$, $\hat{y}_i(\hat{\mathbf{g}}) < \hat{y}_j(\hat{\mathbf{g}})$ and $N_i(\hat{\mathbf{g}}) \setminus \{j\} \subset N_j(\hat{\mathbf{g}}) \setminus \{i\}$ hold. To simplify notation we define the set of agent i and j 's common neighbors, which we denote by $N_{i,j}(\bar{\mathbf{g}}) = \{k \in N \setminus \{i, j\} : \bar{g}_{j,k} = 1 \wedge \bar{g}_{i,k} = 1\}$. Recall that we also defined the following set above $N_{j \setminus i}(\bar{\mathbf{g}}) = \{k \in N \setminus \{i, j\} : \bar{g}_{j,k} = 1 \wedge \bar{g}_{i,k} = 0\}$. Define the following networks $\mathbf{g}' = \hat{\mathbf{g}} - \sum_{k \in N_{j \setminus i}(\hat{\mathbf{g}})} \hat{g}_{j,k}$ and $\mathbf{g}'' = \hat{\mathbf{g}} + \sum_{k \in N_{j \setminus i}(\hat{\mathbf{g}})} \hat{g}_{i,k}$ and the following vectors of effort levels $\mathbf{x}' = (\hat{x}_1(\hat{\mathbf{g}}), \dots, \hat{x}_{j-1}(\hat{\mathbf{g}}), \hat{x}_i(\hat{\mathbf{g}}), \hat{x}_{j+1}(\hat{\mathbf{g}}), \dots, \hat{x}_n(\hat{\mathbf{g}}))$ and $\mathbf{x}'' = (\hat{x}_1(\hat{\mathbf{g}}), \dots, \hat{x}_{i-1}(\hat{\mathbf{g}}), \hat{x}_j(\hat{\mathbf{g}}), \hat{x}_{i+1}(\hat{\mathbf{g}}), \dots, \hat{x}_n(\hat{\mathbf{g}}))$. We will sometimes write $\hat{\mathbf{x}}$ for $\hat{\mathbf{x}}(\hat{\mathbf{g}})$ and \hat{x}_i for $\hat{x}_i(\hat{\mathbf{g}})$. Note that for the network $\hat{\mathbf{g}}$ and the vector of effort levels $\hat{\mathbf{x}}(\hat{\mathbf{g}})$ to be efficient $W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}}) - W(\mathbf{x}', \mathbf{g}') \geq 0$ needs to hold. Therefore, if

$$W(\mathbf{x}'', \mathbf{g}'') - W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}}) - (W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}}) - W(\mathbf{x}', \mathbf{g}')) > 0$$

holds, then $W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}})$ is not efficient. We distinguish two cases. Assume first that $\hat{g}_{i,j} = 0$. We can then write

$$\begin{aligned} & W(\mathbf{x}'', \mathbf{g}'') - W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}}) - (W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}}) - W(\mathbf{x}', \mathbf{g}')) = \\ & \pi_i(\hat{x}_j, \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \hat{x}_k + \sum_{k \in N_{j \setminus i}(\hat{\mathbf{g}})} \hat{x}_k) - \pi_i(\hat{x}_i, \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \hat{x}_k) \\ & - \left(\pi_j(\hat{x}_j, \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \hat{x}_k + \sum_{k \in N_{j \setminus i}(\hat{\mathbf{g}})} \hat{x}_k) - \pi_j(\hat{x}_i, \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \hat{x}_k) \right) \\ & + \sum_{k \in N_{j \setminus i}(\hat{\mathbf{g}})} \pi_k(\hat{x}_k, \sum_{l \in N_k(\hat{\mathbf{g}})} \hat{x}_l + \hat{x}_j) - \sum_{k \in N_{j \setminus i}(\hat{\mathbf{g}})} \pi_k(\hat{x}_k, \sum_{l \in N_k(\hat{\mathbf{g}})} \hat{x}_l) \\ & - \left(\sum_{k \in N_{j \setminus i}(\hat{\mathbf{g}})} \pi_k(\hat{x}_k, \sum_{l \in N_k(\hat{\mathbf{g}})} \hat{x}_l) - \sum_{k \in N_{j \setminus i}(\hat{\mathbf{g}})} \pi_k(\hat{x}_k, \sum_{l \in N_k(\hat{\mathbf{g}})} \hat{x}_l - \hat{x}_j) \right) \\ & + \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \pi_k(\hat{x}_k, \sum_{l \in N_k(\hat{\mathbf{g}})} \hat{x}_l + (\hat{x}_j - \hat{x}_i)) - \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \pi_k(\hat{x}_k, \sum_{l \in N_k(\hat{\mathbf{g}})} \hat{x}_l) \\ & - \left(\sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \pi_k(\hat{x}_k, \sum_{l \in N_k(\hat{\mathbf{g}})} \hat{x}_l) - \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \pi_k(\hat{x}_k, \sum_{l \in N_k(\hat{\mathbf{g}})} \hat{x}_l - (\hat{x}_j - \hat{x}_i)) \right) \end{aligned}$$

Note that linking cost cancel out. We analyze the three differences given above separately. The first difference is zero since $\hat{g}_{i,j} = 0$. If $\frac{\partial^2 \Pi(x,y)}{\partial^2 y} > 0$, then the second and third difference are strictly positive and $W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}})$ is not efficient. If $\frac{\partial^2 \Pi(x,y)}{\partial^2 y} = 0$, then all three differences are zero. Note that for $W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}})$ to be efficient the first order conditions must hold. We know from $\frac{\partial^2 \Pi(x,y)}{\partial x \partial y} > 0$ that not all first order conditions hold in $W(\mathbf{x}'', \mathbf{g}'')$ and there therefore exists a vector of effort levels, $\hat{\mathbf{x}}(\mathbf{g}'')$, such that $W(\hat{\mathbf{x}}(\mathbf{g}''), \mathbf{g}'') > W(\mathbf{x}'', \mathbf{g}'')$.

Next consider the case when $\hat{g}_{i,j} = 1$. The last two differences of the above expression are unaffected and we will suppress them. The effects on i and j in $W(\mathbf{x}'', \mathbf{g}'') - W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}})$ are given by

$$\text{A: } \pi_j(\hat{x}_j, \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \hat{x}_k + \sum_{k \in N_{j \setminus i}(\hat{\mathbf{g}})} \hat{x}_k + \hat{x}_j) - \pi_j(\hat{x}_j, \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \hat{x}_k + \sum_{k \in N_{j \setminus i}(\hat{\mathbf{g}})} \hat{x}_k + \hat{x}_i) \text{ and}$$

$$\text{C: } \pi_i(\hat{x}_j, \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \hat{x}_k + \sum_{k \in N_{j \setminus i}(\hat{\mathbf{g}})} \hat{x}_k + \hat{x}_j) - \pi_i(\hat{x}_i, \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \hat{x}_k + \hat{x}_j),$$

while for $W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}}) - W(\mathbf{x}', \mathbf{g}')$ they are given by

$$\text{D: } \pi_j(\hat{x}_j, \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \hat{x}_k + \sum_{k \in N_{j \setminus i}(\hat{\mathbf{g}})} \hat{x}_k + \hat{x}_i) - \pi_j(\hat{x}_i, \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \hat{x}_k + \hat{x}_i) \text{ and}$$

$$\text{B: } \pi_i(\hat{x}_i, \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \hat{x}_k + \hat{x}_j) - \pi_i(\hat{x}_i, \sum_{k \in N_{i,j}(\hat{\mathbf{g}})} \hat{x}_k + \hat{x}_i).$$

Note that, due to $\frac{\partial^2 \Pi(x,y)}{\partial x \partial y} > 0$ and $\frac{\partial^2 \Pi(x,y)}{\partial^2 y} \geq 0$, the expression in A is larger than the expression in B and the expression in C is larger than in D. That is, again $W(\mathbf{x}'', \mathbf{g}'') - W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}}) - (W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}}) - W(\mathbf{x}', \mathbf{g}')) > 0$ holds and $W(\hat{\mathbf{x}}(\hat{\mathbf{g}}), \hat{\mathbf{g}})$ is not efficient.

Part (iii) For $\Phi(\cdot)$ strictly convex the conditions $\Phi(1) \simeq 0$ and $\Phi(\frac{n(n-1)}{2}) > \sum_{i \in N} \pi_i(\hat{\mathbf{x}}(\bar{\mathbf{g}}^c), \bar{\mathbf{g}}^c)$ guarantee that the complete and the empty network are not part of any efficient configuration and the efficient network is therefore a nested split graph that is neither complete, nor empty. *Q.E.D.*

9 References

1. Adamic, L.A. and E. Adam, (2003), Friends and neighbors on the web, *Social Networks* 25, 211-230.
2. Ahuja, G., (2000), Collaboration networks, structural holes and innovation: A longitudinal study, *Administrative Science Quarterly*, 45, 425-455.
3. Åkerman, A. and A. Larsson (2014), The global arms trade network 1950–2007, *Journal of Comparative Economics*, 42 (3), 535-551.
4. Allouch, N., (2015), On the private provision of public goods on networks, *Journal of Economic Theory*, 157, 527–552.
5. Baker, G., R. Gibbons, and K.J. Murphy, (2004), Strategic alliances: bridges between islands of conscious power, mimeo, MIT.
6. Bala, V. and S. Goyal, (2000), A non-cooperative model of network formation, *Econometrica*, 68 (5), 1181-1229.
7. Baldwin, T.T., M.D. Bedell, and J.L. Johnson, (1997), The social fabric of a team-based M.B.A. program: network effects on student satisfaction and performance, *Academy of Management Journal*, 40, 1369-1397.
8. Ballester, C., A. Calvó-Armengol, and Y. Zenou, (2006), Who is Who in Networks. Wanted: The Key Player, *Econometrica*, 74, 1403-1417.
9. Bayer, P., Hjalmarsson, R., D. Pozen, (2009), Building criminal capital behind bars: Peer effects in juvenile corrections, *The Quarterly Journal of Economics*, 123 (1), 105-147.
10. Bätz, O., (2015), Social Activity and Network Formation, *Theoretical Economics*, 10 (2015), 315–340.
11. Belhaj, M., Bervoets, S. and Deroïan, F., (2016), Efficient Networks in Games with Local Complementarities, *Theoretical Economics*, 11 (3), 357-380.
12. Belhaj, M., Bramoullé, Y. and Deroïan, F. (2014), Network games under strategic complementarities, *Games and Economic Behavior*, 88, 310–319
13. Belleflamme, P., and F. Bloch, (2004), Market Sharing Agreements and Stable Collusive Networks, *International Economic Review*, 45 2, 387-411.
14. Bramoullé, Y., (2007), Anti-Coordination and Social Interactions, *Games and Economic Behavior*, 58, 30-49.
15. Bramoullé, Y. and D’Amours, M, (2014), Strategic interaction and networks, *American Economic Review*, 104 (3), 898-930.
16. Cabrales, A., Calvó-Armengol, A. and Y. Zenou (2011), Social interactions and spillovers, *Games and Economic Behavior*, 72, 339-360.

17. Calvó-Armengol, A., E. Patacchini and Y. Zenou, (2009), Peer Effects and Social Networks in Education, *Review of Economic Studies*, 76, 1239-1267.
18. Calvó-Armengol, A., Patacchini, E. and Y. Zenou (2005), Peer effects and social networks in education and crime, *CEPR Discussion Paper* No. 5244.
19. Canter, D. (2004), A Partial Order Scalogram Analysis of Criminal Network Structures, *Behaviormetrika*, 31, No.2, 131-152.
20. Chvátal, V. and P. Hammer (1977), Aggregation of inequalities in integer programming, *Annals of Discrete Mathematics* 1, 145 -162.
21. Case, A. C. and L. F. Katz, (1991), The Company you Keep: The Effects of Family and Neighborhood on Disadvantaged Youths, *NBER Working Paper* No. 3705.
22. Cohen, W. and D. Levinthal, (1989), Innovation and learning: The two faces of R&D, *The Economic Journal*, 99, 569-596.
23. Cohen, W. and D. Levinthal (1990), Absorptive capacity: A new perspective on learning and innovation, *Administration Science Quarterly*, 35, 128-152
24. Cohen-Cole, E., E. Patacchini, Y. Zenou (2012), Systemic Risk and Network Formation in the Interbank Market, mimeo.
25. Conley, T. G. and G. Topa, (2002), Socio-Economic Distance and Spatial Patterns in Unemployment, *Journal of Applied Econometrics*, 17, Issue 4, 303-327.
26. Corbae, D., M. Stinchcombe and J. Zeman, (2009), An Introduction to Mathematical Analysis for Economic Theory and Econometrics, *Princeton University Press*.
27. Damm, A. P. and Dustmann, C. (2014), Does growing up in a high crime neighborhood affect youth criminal behavior?, *American Economic Review*, 104 (6), 1806-1832.
28. De Benedictis, L., L. Tajoli, (2011), The world trade network, *The World Economy* 34 (8), 1317-1454.
29. de la Fuente, A. (2009), *Mathematical Methods and Models for Economists*, Cambridge University Press.
30. Dutta, B., Ghosal, S. and D. Ray (2005), Farsighted network formation, *Journal of Economic Theory*, 122, 143-164.
31. Galeotti, A. and S. Goyal, (2010), The Law of the Few, *American Economic Review*, 100, No. 4, 1468-92.
32. Glaeser, E., B. Sacerdote and J. Scheinkman (1996), Crime and Social Interactions, *The Quarterly Journal of Economics*, 111, No. 2, 507-48.
33. Goyal, S. (2007), *Connections: An Introduction to the Economics of Networks*, Princeton University Press.

34. Goyal, S., and S. Joshi (2003), Networks of Collaboration in Oligopoly, *Games and Economic Behavior* 43, 57-85.
35. Hagberg A., P. Swart, D. Schult, (2006), Designing threshold networks with given structural and dynamical properties, *Physical Review E*, 74, 056116.
36. Hiller, T., (2013), Peer Effects in Endogenous Networks, <http://sticerd.lse.ac.uk/dps/te/te564.pdf>
37. Hoxby, C., (2000), Peer Effects in the Classroom: Learning from Gender and Race Variation, *NBER Working Paper* 7867.
38. Kinatered, M. and Merlino L. P., (forthcoming), Public Goods in Endogenous Networks, *American Economic Journal: Microeconomics*.
39. Jackson, M.O., and J. Wolinsky, (1996), A Strategic Model of Social and Economic Networks, *Journal of Economic Theory*, 71 (1), 44-74.
40. Kennan, J., (2001), Uniqueness of Positive Fixed Points for Increasing, Concave Functions on \mathbb{R}^n : An Elementary Result, *Review of Economic Dynamics*, 4, 893-899.
41. König, M., C. J. Tessone, and Y. Zenou, (2014), Nestedness in Networks: A Theoretical Model and Some Applications, *Theoretical Economics*, 9 (3), 695-752.
42. Lagerås, A. N., Seim, D., (2016), Strategic complementarities, network games and endogenous network formation, *International Journal of Game Theory*, 45.3, 497-509.
43. Levin, R. C. and Reiss, P. C. (1988), Cost-reducing and demand-creating R& D with spillovers, *Rand Journal of Economics*, 19, 538-56.
44. Lin, M., (2001), Social Capital: A Theory of Social Structure and Action, Cambridge University Press, Cambridge.
45. Lovasz, L., (2009), Geometric Representations of Graphs, <http://www.cs.elte.hu/~lovasz/geomrep.pdf>
46. Ludwig, J., Duncan, G.J. and Hirschfield, P. (2001), Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment, *The Quarterly Journal of Economics*, 116 (2), 655-679.
47. Powell, W.W., K.W. Koput, L. Smith-Doerr and J. Owen-Smith, (1999), Network position and firm performance: organizational returns to collaboration in the biotechnology industry, in: S. Andrews, D. Knoke (Eds.), *Networks in and Around Organizations*, JAI Press, Greenwich, 129-160.
48. Sacerdote, B., (2001), Peer Effects with Random Assignment: Results for Dartmouth Roommates, *Quarterly Journal of Economics*, 116, No. 2, 681-704.
49. Soramaki, K., M. Bech, J. Arnold, R. Glass, W. Beyeler, (2007), The topology of interbank payment flows, *Physica A: Statistical Mechanics and its Applications* 379 (1), 317-333.

50. Topa, G., (2001), Social Interactions, Local Spillovers and Unemployment, *Review of Economic Studies*, 68 (2), 261-95.
51. Uzzi, B., (1996), The sources and consequences of embeddedness for the economic performance of organizations: The network effect, *American Sociological Review*, 61, 674–698.